Investigating the effect of a quality improvement intervention on newborn care in three central districts in Malawi

Bejoy Padmanabhan Nambiar
Institute for Global Health
University College London

Supervisor : Dr.Sarah Hawkes

Co-Supervisor : Prof.Anthony Costello

Dr. Timothy Colbourn

Thesis submitted for the degree of Doctor of Philosophy
January, 2016

Abstract

Quality Improvement (QI) interventions are complex in nature and evaluation focusing solely on measuring its impact is unable to answer questions of how or why the intervention worked. Results from the impact and outcome evaluation of a large QI intervention aimed at reducing newborn mortality in 3 central districts in Malawi had been inconclusive. I use a theory based evaluation (TBE) approach in my PhD, to analyse the implementation theory and understand the program theory behind the QI intervention in Malawi.

I developed a research strategy, which includes a combination of Theory of Change (ToC) and Realist evaluation approaches for program evaluation. I begin with a description of the Theory of Change and how it evolved over the course of the project. Following this I produce a summary and synthesis of all the available evidence in relation to this Theory of Change. This includes all data generated from the process evaluation as well as a range of documentary evidence generated through the course of the project. For synthesizing this evidence I use the consolidated framework for implementation research (CFIR), which helps to develop a structured understanding of the implementation theory. Next, I adapt realist evaluation principles to analyse the intervention mechanisms. I compare the mechanisms of successful QI interventions with those of the Malawi intervention to understand why we were unable to observe an impact on newborn care in our QI health facilities.

Various mechanisms for successful QI interventions have been hypothesized in literature. The strategy adopted to trigger these mechanisms include collaborative learning sessions, data improvement, continuous stakeholder engagement and feedback. While all these strategies could be identified within the QI intervention implemented in Malawi the implementation process was unable to garner enough individual or organizational momentum to trigger the actual intervention mechanisms. This can be attributed to a series of implementation challenges such as the intervention complexity characterized by conceptual clarity regarding the intervention design, the evolving role of the implementing organization, stakeholder perceptions, programme planning and governance. Contextual factors such as availability of human and material resources also influenced the mechanism of the intervention.

Impact evaluation of complex interventions is important, but seldom provides information to decision makers and policy advisors on 'how' interventions work. Using theory-based approaches to evaluating complex QI interventions provides an insight into how and why the interventions worked (or not). Such evaluation is important for theory building and needs to adopt a broader health systems research (HSR) approach to evaluation In the discussion, I critique the evaluation approaches used and provide recommendations for future evaluation of complex interventions.

Acknowledgements

My journey through this thesis has been a fantastic learning exercise. This would not have been possible without the support of my PhD Supervisors. Firstly, I would like to thank Dr. Sarah Hawkes for agreeing to be my Supervisor and having the patience and perseverance to nudge me to the finish line. I would like to thank Dr. Tim Colbourn, more a friend than a supervisor, for his unrelenting support and encouragement. A special thanks to Prof.Costello for believing in me and giving me the opportunity to work with his wonderful team at IGH.

I would like to express gratitude to my family for their unconditional love and support and for endeavouring with me through this journey. No words are enough to thank my wife Mahima for living up with my eccentricities and taking on all the additional responsibilities with a smile. It is going to be a tall order returning back all those favours in the post-PhD era. I would also like to take this opportunity to thank my parents, parents-in-law, my sister and her husband, without whose support I would not have been able to make this journey.

I would like to specially thank my dear friend, Dr.Katharina Kranzer for her selfless support and motivation that she has provided throughout my PhD and for hosting me during my numerous trips to London.

I would not have been able to reach this far in my PhD without the support of friends and colleagues at MaiKhanda and PACHI, the organizations that I have been involved with in Malawi since the day I arrived. I would like to offer a special thanks to Dr. Peter Kazembe for providing office space in my early days in Malawi and to Dr. Charles Mwansambo for being there to support ideas and initiatives.

A sincere word of thanks to all the mothers and babies as well as staff in the health facilities who were involved in the original MaiKhanda project.

Declaration

I, Bejoy P.Nambiar, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm this has been indicated in the thesis.

Preface

This thesis is organized as follows. The introduction, Chapter 1, provides a brief introduction to the MaiKhanda project and the Quality Improvement intervention that it had initiated in health facilities across 3 central districts in Malawi. The Chapter also provides an overview of the evaluation design that was adapted at MaiKhanda to measure impact and process of the intervention. Chapter 2 begins with a clarification of the concept of quality of care and quality improvement followed by the different approaches currently used in the practice of quality improvement. The next section in this Chapter is a literature review of the quality improvement approach that was used in MaiKhanda-the Quality Improvement (QI) Collaborative. Chapter 3 presents a summary of the impact evaluation. It also explains the limitations of a RCT (Randomized Control Trial) design in evaluation complex intervention such as QI. This lays the foundation for alternate theory based approaches to explain the findings of the impact evaluation. Chapter 4 presents the research strategy that I have adopted in my thesis to analyse the program theory and implementation theory of MaiKhanda interventions. Chapter 5 describes the evolution of MaiKhanda's Theory of Change over the course of the project. Chapter 6 summarizes all the available evidence generated from MaiKhanda to produce a cohesive picture of intervention progress through the project time period. The evidence repository includes the process evaluation studies conducted by the UCL evaluation team from 2007-2011, as well as all the project documents such as bi-annual and quarterly progress reports and programme meeting minutes. Chapter 7 is the analysis chapter and it provides an explanation of MaiKhanda's program theory as well as implementation theory. Chapter 8 discusses the key findings from the study and highlights some of the methodological issues in conducting evaluation using theory based evaluation approaches. It also discusses some of the key

learning for future research. The last section of this Chapter summarise the limitations in my study approach. Chapter 9 concludes with a summary of key findings as well as issues and challenges in evaluating QI in resource poor settings and suggests a whole systems improvement approach as the way forward.

The appendices contain a summary of the all the relevant process evaluation studies conducted as part of the MaiKhanda evaluation and used for secondary analysis in my PhD. This describes the various constructs considered in my evaluation and their evolution over the different intervention phases. As part of the requirement towards my UCL doctorate, I have attached as Appendix-1, a summary of my training and professional development, including courses and conference attended and fieldwork.

This PhD makes methodological contributions by identifying a novel research strategy to evaluate complex QI interventions. The research strategy combines a probability and plausibility design to answer questions related to impact as well as understanding the mechanism of how social interventions such as QI work. The PhD also explores literature surrounding the evaluation of quality improvement interventions. It hopes to contribute towards improving the quality of newborn care in Malawi and more broadly, towards developing an evidence based approach to quality improvement interventions.

This PhD is a secondary analysis of the data collected as part of the MaiKhanda QI evaluation. I had access to MaiKhanda resources, including trained Monitoring and Evaluation Officers to collect and collate the data for the various process evaluation studies considered in my PhD. The work presented in this thesis, however, is entirely my own, as are any mistakes.

Bejoy Nambiar,
Lilongwe, Malawi
January, 2016

Ethics

The original MaiKhanda study had ethical approval from the National Health Sciences Research Committee in Malawi and UCL ethics committee. My PhD was a secondary analysis of the data collected as part of UCL engagement in evaluation of the MaiKhanda project and as such did not warrant a separate ethical clearance.

Funding

Funding was provided by The Health Foundation, which paid for my salary as a full time Research Associate at the Institute for Global Health. Funding was from the period January, 2007 to February, 2012. Subsequently, I have been supported with funding from European Union, Bill and Melinda Gates Foundation and Department for International Development (UK).

Table of Contents

List of Figures

List of Tables

Glossary of Terms

ALSO	Advanced Life Support Obstetric Course
CBA	Criterion Based Audit
CHAM	Christian Health Association of Malawi
DHS	Demographic and Health Survey
BEMOC	Basic Emergency Obstetric Care
CBDA	Community Based Distribution Agent
CEMOC	Comprehensive Emergency Obstetric Care
DHMT	District Health Management Team
DHO	District Health Office
GFATM	Global Fund for AIDS ,Tuberculosis and Malaria
HMIS	Health Management Information System
HSA	Health Surveillance Assistants
HSR	Health Systems Research
IHI	Institute for Healthcare Improvement
IMCI	Integrated Management of Childhood Illnesses
JHPIEGO	Johns Hopkins Program for International Education in Gynaecology and Obstetrics
KI	Key Informant
KMC	Kangaroo Mother Care
LSTM	Liverpool School of Tropical Medicine

LATH	Liverpool Associates in Tropical Health
MD	Maternal Death
MDR	Maternal Death Review
MEO	Monitoring and Evaluation Officers
MICS	Multiple Indicator Cluster Survey
MoH	Ministry of Health
MSCE	Malawi School Certificate of Education
MSH	Management Science for Health
PDSA	Plan Do Study Act (Cycles)
PVA	Perinatal Verbal Autopsy
QI	Quality Improvement
QIC	Quality Improvement Collaborative
RHU	Reproductive Health Unit
RMNCH	Reproductive Maternal Newborn and Child Health
STTA	Short Term Technical Assistance
THFC	The Health Foundation Consortium
TBE	Theory Based Evaluation
TBIE	Theory Based Impact Evaluation
US	United States of America
UK	United Kingdom
WHO	World Health Organization

Chapter 1 Introduction

1.1 Maternal and Newborn Health

Approximately 300,000 women are estimated to die of pregnancy-related causes worldwide, of which 99% occurs in the developing world(1, 2). More than 50% of the countries with increased maternal mortality rates are in Sub Saharan Africa. However, there are affordable and effective interventions that can prevent or avoid more than 50% of these deaths(3)

An estimated 2.76 million newborn deaths occur each year, accounting for 44% of under-five deaths(4). Low cost interventions with proven efficacy are available to improve newborn survival(5). As we move on from Millennium Development Goals to an era of Sustainable Development Goals, one of the key factors influencing and sustaining improvements in maternal and newborn care is quality. Quality has been described as the third revolution in global health(6).

1.2 Maternal and newborn health in Malawi

With a GDP of $780 per capita, Malawi (Figure 1) is one of the poorest countries in Sub-Saharan Africa(7). It is landlocked and mainly dependent on agriculture, especially tobacco, for its income (contributing to 35% of the GDP and 80% of the employment). Malawi has a population of 17.2 million with an annual population growth rate of 2.8%(7). Most people (85%) live in rural areas. The crude birth rate is 39.2 per 1000 population per year while the crude death rate is 8.6 per 1,000 population per year. With an HIV prevalence of 11%, Malawi is one of the countries severely affected by the AIDS epidemic. The Total Fertility Rate has been on the decline, with the current TFR being 5.0 overall, and lower in urban areas. The Maternal mortality rate has been estimated at 574 per 100,000 live births and the newborn mortality rate is 29 per 1,000 live births(8)

The total health care spending for Malawi in 2012 was $625 million, which amounted to a health care spending of $39.3 per capita. The total spending

on health in 2012 was 9% of the GDP[1]. Donor investments in general increased over the years. In 2009, donors contributed about 62% of the total funds and within MNCH (Maternal Newborn and Child Health) it was about $5.3million(9, 10).

Figure 1: Map of Malawi

Although recent trends show a decline in mortality, the maternal mortality in Malawi is high. The main direct causes of maternal death in Malawi are haemorrhage, sepsis, eclampsia, and ruptured uterus; the main indirect causes are anaemia and HIV. The uptake of antenatal care is very high at 96% for at least 1 antenatal visit, although women undergoing the recommended 4 visits is lower at 44%. Institutional deliveries are at 89% as of 2014 and the uptake of post-natal care is 75%(8).

Infant mortality is at 53 per 1000 live births, and neonatal mortality at 29 per 1000 live births in Malawi. While there has been a steady decline in under five mortality at an annual average rate of 5.4%, over the past decade, the decline

has been relatively smaller for newborn health. While Malawi will be well within the MDG targets for child mortality, further improvements in this area will need evidence based investments in newborn care(10). The main causes of neonatal death are prematurity and low birth weight, asphyxia and sepsis(11).

There is a high political commitment for maternal and newborn health in Malawi(12). This has provided a programmatic and policy platform to address quality gaps in health facilities as well as initiate community based newborn care. In early 2005, the Reproductive Health Unit (RHU) at the Ministry of Health developed a Road Map for reducing maternal and infant mortality. The general recommendations indicated a need to improve the quality of maternal and newborn health care in both the health service delivery system (supply side) and the community (demand side).

It was under these circumstances that a combined supply and demand side intervention addressing quality gaps in facilities and mobilizing women in the communities respectively, was set up and what came to be known as the MaiKhanda programme. The objective of the programme was to impact on maternal and newborn survival in Malawi.

1.3 Background

The Health Foundation (a UK based charity) with a history of working in QI (Quality Improvement) in the NHS (National Health Services, UK), invited partners IHI (Institute for Healthcare Improvement), LSTM (Liverpool School of Tropical Medicine), WCF (Women and Children First-a UK based NGO) to form a Consortium to establish a programme in three districts in central Malawi, which came to be known as the MaiKhanda (Chichewa for Mother [Mai] and Neonate [Khanda]) programme.

The MaiKhanda programme started its initial activities in Malawi in February 2006 through a stakeholder meeting between The Health Foundation Consortium (THFC) members and key stakeholders from the MoH (Ministry of Health) and health facilities, explaining the purpose of the programme.

The Malawi implementation team recruitment started in February 2006 and by June 2006 MaiKhanda began work initially in 4 and then gradually 9 CEmOC[i] facilities in the three Malawian districts (Kasungu, Lilongwe and Salima) with a Quality Improvement (QI) intervention. The QI intervention worked with providers to improve health care practices with the overall aim to reduce maternal and neonatal mortality.

The programme was innovative, in that it combined 2 large-scale interventions addressing both, the supply and demand side of maternal and newborn care:

- a community mobilization intervention working primarily through women's groups using a participatory action cycle, that had shown some promising evidence from international studies.
- a quality improvement intervention consisting of Learning Collaboratives and Action Periods in the health facilities, working primarily through QI teams in the health facilities.

The programme was implemented in the 3 central region districts of Salima, Kasungu and Lilongwe over a 6 year period and mobilized a total of 879 women's group and worked with 9 hospitals and 29 health centres across the 3 districts to identify and implement local strategies for maternal and newborn health improvement.

The Health Foundation commissioned Institute for Global Health (IGH) at University College London (UCL) to undertake an independent evaluation of the programme between 2006 and 2010. The primary outcomes looked at the effect of the two interventions, individually and in combination, on maternal, neonatal and perinatal mortality. In addition, the study also included a process evaluation and an economic evaluation for the proposed interventions.

---

[i] *CEmOC refers to Comprehensive Emergency Obstetric Care. "Basic" and "Comprehensive" emergency obstetric and care (EmOC) are defined by their signal functions: A Basic EmOC facility (BEmOC) should offer skilled attendance at birth, administration of IV fluids, antibiotics, anticonvulsants and oxytocics, manual removal of the placenta, provide help for retained products of conception (e.g. manual vacuum aspiration MVA) and carry out an assisted delivery when necessary. A Comprehensive EmONC facility (CEmONC) should additionally offer blood transfusions and caesarean sections.*

1.4 MaiKhanda's Theoretical Framework

The programmes' theoretical framework (Figure 2) is based on the three delays model by Thaddeus and Maine(13) where delay 1 is defined as the decision to seek care, delay 2 is the delay in getting to a health facility in time and the third delay is the delay in getting adequate and appropriate treatment at the health facility.

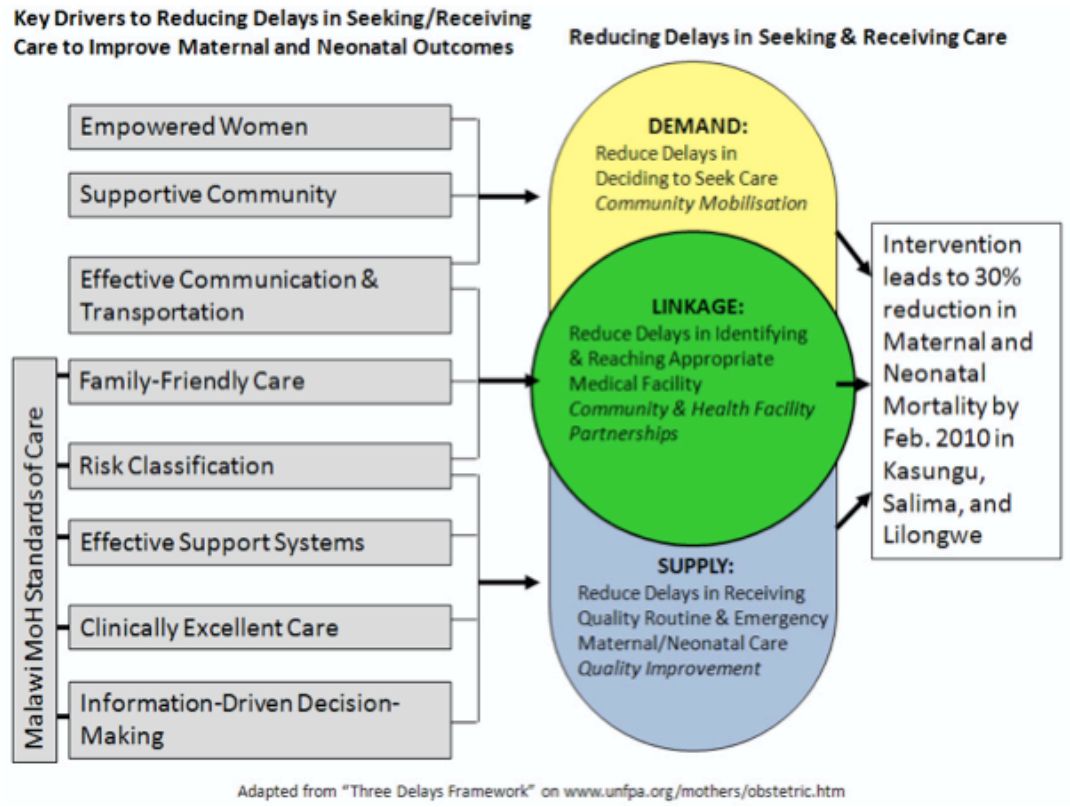

Figure 2: MaiKhanda's theoretical framework

The framework envisaged a single programme of work where communities mobilized through women's group interventions would impact on delay 1 and delay 2, while using a quality improvement intervention would reduce delay 3. Using this approach, the programme hypothesized that it would reduce maternal and neonatal mortality in the selected 3 districts in the central region of Malawi over a 5 year period. The interventions are described in detail below.

1.5 Intervention Design

1.5.1 Demand-side (Community) intervention

The demand side interventions (community interventions) were focused on delays 1 and 2. The intervention consisted of the women's group action cycle where women from the community were brought together regularly (usually monthly) to discuss issues around pregnancy and birthing and to find solutions to their problems. They would go through an action cycle, which had four phases. Phases I and II consisted of a series of 8 meetings to identify and prioritize issues around maternal and newborn health; these were followed by strategy development (Phase III) and evaluation (Phase IV). A facilitator who is usually a community member facilitated these action cycles. Several models for the facilitators had been tried previously - In the MIRA (Nepal) trial(14) and a recently finished trial in Mchinji district(15) in Malawi the women's group facilitators were paid employees of the project, while a similar women's group intervention in Ntcheu district in Malawi has used health surveillance assistants (HSAs) from the existing health system to facilitate the women's group interventions. MaiKhanda had women volunteering from the community acting as women's group facilitators.

1.5.2 Supply-side facility intervention

The supply side intervention[ii] consisted of a model, which integrated Rapid Cycle process improvement (Plan-Do-Study-Act, PDSA) with Criterion-Based Audit (or standard-based audit) cycle. This integrated model, along with Maternal (and Neonatal) Death Reviews formed the components of MaiKhanda's 'quality improvement' intervention in the facilities. This integrated model for improvement was implemented through a process called 'Breakthrough Collaborative Series' which included collaborative Learning Sessions amongst Quality Improvement (QI) teams from different health facilities and Action Periods where individual QI teams received mentoring and support from MaiKhanda to implement PDSA cycles in their respective

[ii] *The focus of my PhD was on understanding the effect of QI interventions on newborn mortality. My thesis does not involve an analysis of the demand side intervention.*

facilities (Figure 3). The choice of topic for improvement was guided by the newborn change package. The change package for newborn care included neonatal resuscitation; kangaroo mother care and sepsis management. During Learning Session, health facilities would come together and share their learning and experiences with each other about their improvement work, and during the Action Period MaiKhanda QI officers would make field visits to health facilities and facilitate implementation of their proposed changes, as agreed during the learning session. Each Learning Session was conducted over a 90 day period and Action Period constituted the time in between the two learning sessions. The QI Collaborative is discussed in detail in Chapter 2 (§2.3.4).

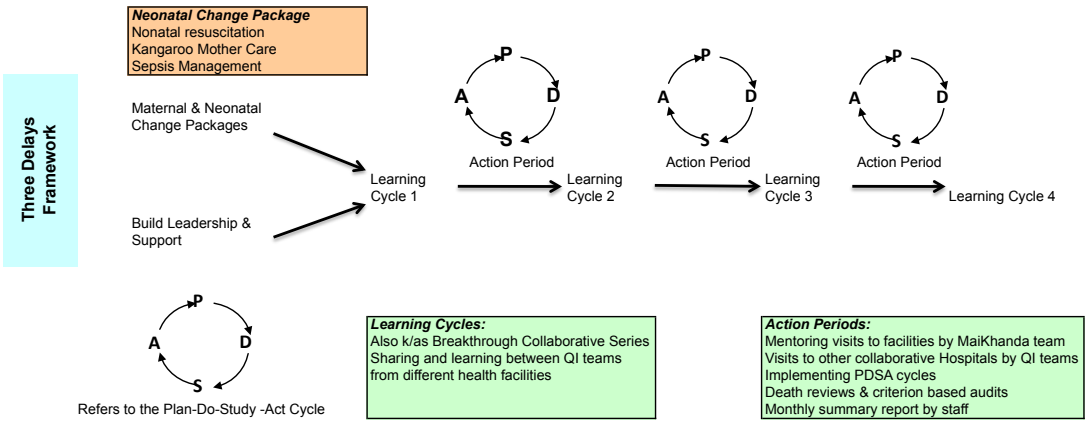

Figure 3: MaiKhanda's improvement model

The Criterion Based Audit (CBA) approach was introduced and carried out by health care providers at the facility level, to achieve changes in practice and improve quality of care. It involved a process of reviewing existing guidelines, and was linked with the standards of practice. Standards of practice were developed by Ministry of Health to act as a guide for providers regarding the best standards of care in specific areas of healthcare. The CBA cycle is

repeated as often as needed to improve the level at which standard is achieved.

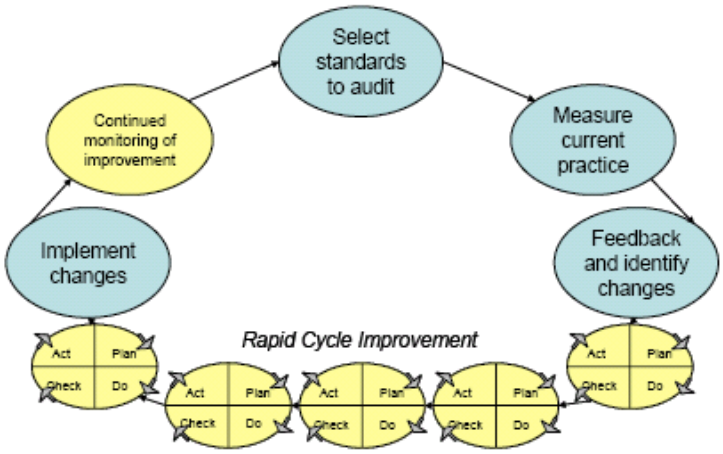

Figure 4: Integrated Model for Improvement

The *Rapid Cycle (PDSA) approach* also built on the providers' knowledge base and was to be done through a structured series of small interventions, designed by the providers themselves, and focussed on a more efficient use of existing resources. The integration of the two approaches (Figure 4) would occur with the process beginning with an audit cycle and then moving into rapid cycle change when the problems had been identified. This model used a 'bottom-up', participatory approach and was in line with recommendations from the maternal and newborn health roadmap developed by RHU (Reproductive Health Unit) in 2005.

Ministry of Health recommendations also included the implementation of Death Reviews. MaiKhanda facilitated the local staff to conduct audits in a blame-free and participatory manner. The information generated by the review would then be used to determine the focus of the criterion-based and rapid-cycle improvement activities.

This integrated model was reviewed towards the middle of the programme and was revised to adapt a Model for Improvement focussed more on the Collaborative Learning and Action Cycles (§5.1.3).

1.6 Intervention timeline

The intervention was initially planned over a five-year period (2006-2010) covering three districts. In phase one, which would last three years (2006-08), facility improvement was to reach two-thirds of all facilities in the districts and phase two (2008-2010) of the programme was to cover all the facilities in the districts. During phase two of the intervention, the 'dosage' of the facilities intervention was greatly increased from Oct-09 onwards, with more than 60% of the planned activities being completed. Therefore, when exploring possible differences in intervention effect over time, the facilities intervention period has been further split into two time periods. Thus the intervention can be considered of the following 4 phases (Table 1):

Table 1: MaiKhanda Intervention Phases

Intervention Phases	Time Period
Baseline period	June, 2006 to January, 2007
Pre-intervention period	February, 2007 to September, 2008
Early intervention period	October, 2008 to September, 2009
Late Intervention period	October, 2009 to December, 2010

A summary of events in relation to the supply side (QI) intervention is outlined in Table 2

Timeline	Event
January 2006	Proposal approved by The Health Foundation
	Quality Improvement pilot in Bwaila hospital
July, 2006	Quality Improvement initiated in CEmOC facilities
	Maternal Death Reviews initiated
January, 2007	First QI breakthrough collaborative
	Introduction of Criterion-Based Audits
February, 2007	1 QI officer appointed
July 2007	Surveillance system for impact evaluation established
July 2007	Health Centre work initiated by LSTM
	Programme Director recruited
Sept, 2007-	MaiKhanda strategic review and planning
Feb, 2008	LSTM and LATH leave Consortium
	THF grants an additional year to the project for Phase II
March 2008	Project gets registered as an NGO in Malawi-MaiKhanda
May 2008	Second round of proposals with a clear design in place
June, 2008	Super-Improver training
July, 2008	Health centre work re-initiated by IHI
Feb, 2009	Implementation of neonatal change ideas
July, 2009	Health Centre work being fully implemented
Sept, 2009	FI Officers based in Salima and Kasungu districts and extra officers in Lilongwe district
2010	Bi-weekly visits to health centres; regular visits to CEmOCs

Table 2: MaiKhanda QI intervention timeline

1.7 Evidence base for MaiKhanda's interventions

Several prominent studies suggested the significant impact of community level interventions on maternal and neonatal mortality. A meta-analysis of 7 trials using women's group participatory learning and action cycles showed a 37% reduction in maternal mortality (OR 0.63, 95% CI 0.32-0.94) and a 23% reduction in newborn mortality (OR 0.77, 95% CI 0.65-0.90) (16).

Thus evidence regarding women's group participatory action cycles for maternal and newborn health is well established. However, evidence regarding the effect of improvement models on maternal and newborn health is very limited and especially for low-income countries it is practically non-existent.

The evidence for the model for quality improvement practiced at MaiKhanda was based on IHI's (Institute for Healthcare Improvement) experience of conducting quality improvement work in the US and other high income countries such as UK and Europe. Health care providers in market driven health care settings such as the United States need to demonstrate high-quality care to compete in an oversupplied health care market. While in publicly funded health systems such as Malawi, health care providers need to improve quality to make scarce resources stretch further(17). Therefore, the intentions and objective of various health care systems in improving quality has an effect on the way the project is delivered at the ground level. Furthermore, there was no evidence to show that the methods used in high-income settings were appropriate and could produce similar results in low-income countries as well.

Nevertheless, even with the limited evidence available, most of the studies did not report on their primary outcomes. Franco and Marquez detailed non-controlled time-series evidence in support of quality improvement breakthrough series Collaboratives leading to improvement in adherence to essential standards of care and some health outcomes in 12 low and middle income countries(18). A phased, rapid national scale up quality improvement intervention focused on under five children in Ghana, tested process changes

using QI techniques, along the continuum of care from pregnancy to age 5 in both inpatient and outpatient settings but failed to report on the impact of the intervention(19). A systematic review by Schouten and colleagues on the evidence for quality improvement Collaboratives identified 57 studies (from a total of 72 studies) that were based on the breakthrough series Collaboratives. A majority of these studies did not have a comparison group for impact assessment(20). The USAID sponsored Health Care Improvement (HCI) Project in Niger focused on institutionalization of quality improvement for essential obstetric and neonatal care and introduced change packages across different levels of their healthcare system. The assessment looked mainly at performance measures but did not report on primary outcomes (21, 22).

The MaiKhanda intervention was the first of its kind to measure the impact of QI intervention within the context of a low-income setting such as Malawi. The next Section presents an overview of the evaluation design used in the study.

1.8 Evaluation design overview

Based on the Donabedian framework of input-process and output(23), the evaluation was designed to include an impact evaluation, process evaluation and economic evaluation (Figure 5).

1.8.1 Impact Evaluation

The interventions were evaluated via a two-by-two factorial cluster randomised controlled trial design, where all pregnant women in surveillance areas were eligible to take part and consenting women were followed-up to two months after birth via a low-cost community surveillance system using village-based key informants. Primary outcomes were maternal, perinatal and neonatal mortality. There were a total of 61 health centre population clusters evaluated in the trial.

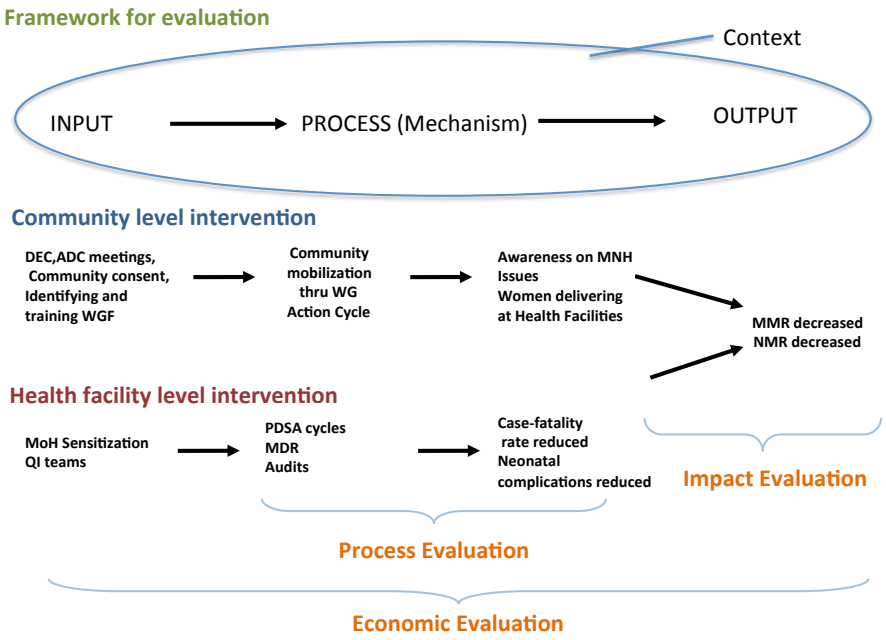

Figure 5: MaiKhanda Evaluation Framework

The two-by-two factorial design included health centres (clusters) randomised to either receive or not receive the Quality Improvement (QI) intervention and the communities surrounding each group of health centres then stratified and randomised to either receive or not receive the Women's Group (WG) intervention(24). Thus each health facility and its catchment population were randomly allocated to receive a women's group or quality improvement intervention or both or none (control group). Randomisation was stratified by the two interventions and by district, so that the numbers of intervention and control clusters in each district balanced.

A separate time-series evaluation of the quality improvement work in 9 CEmOC hospitals, that were not part of the trial design, was under-taken with maternal and neonatal case-fatality rates as primary outcomes. These 9 CEmOC facilities were not included in the randomization process as they received referrals from both intervention and control health centres and including them would lead to 'contamination' of the groups. Nevertheless, QI interventions were implemented in these facilities as part of the intervention

design to implement QI across a 'vertical slice' of the health system namely primary, secondary and tertiary care facilities.

Details of impact evaluation implementation and trial results are discussed in detail §Chapter 3.

1.8.1.1 Secondary outcomes

In order to track the secondary outcomes in health centre and hospital and to compliment the population level mortality data collected as part of the RCT, the evaluation also undertook monthly health facility surveillance. The secondary outcomes included the signal functions as well as data on maternal and neonatal deaths and case-fatality rates. These were collected on a monthly basis directly from the Health Centres and from the Hospitals (CEmOC) using a standard Health Facility Surveillance form. Trends in facility case-fatality rates were assessed by time series analysis and quality improvement control chart statistics.

1.8.2 Process evaluation of QI at health facility level

Evaluations focusing exclusively on outcomes face a number of limitations and are unable to answer a number of important questions, since they do not explore the contextual factors, the design and delivery of the intervention, the extent and quality of the delivery of the intervention, the mechanism of the intervention and the differential effects of the intervention on different populations(25). Without this information it is difficult to explain why an intervention might have worked (or failed to work), to ensure the replication of interventions in other settings or to assist in the further improvement of effective interventions.

More recently evaluators has begun to take cognizance of the fact that these limitations to experimental study designs can be overcome to a large extent by adequately describing and monitoring an intervention and its implementation through a process evaluation along side conventional evaluation methods(26). Thus process evaluation has emerged as an important component of randomised controlled trials (RCTs) of complex public

health interventions(27). In integrating process evaluation with RCTs, the evaluation offers researchers, decision makers and policy implementers a much deeper understanding of the interventions for replication and scaling up(28). Recent trials of complex interventions have integrated process evaluation to understand and explain the trial outcomes(29). It has been recommended as a preferred method especially to address the gaps in RCTs for evaluating social interventions (30).

MaiKhanda's evaluation design consisted of a process evaluation to look into the 'black box' of the intervention and describe the key lessons learnt, identify the scope for future improvement and outline the implications for policy.

Based on the quality of care framework suggested by Hulton(31) for maternal and newborn health care, the evaluation design considered three determinants of the quality of service provision. These were availability of human and material resources, provision of care and patient experience of care as well as the macro-system of policies within the local and national health system.

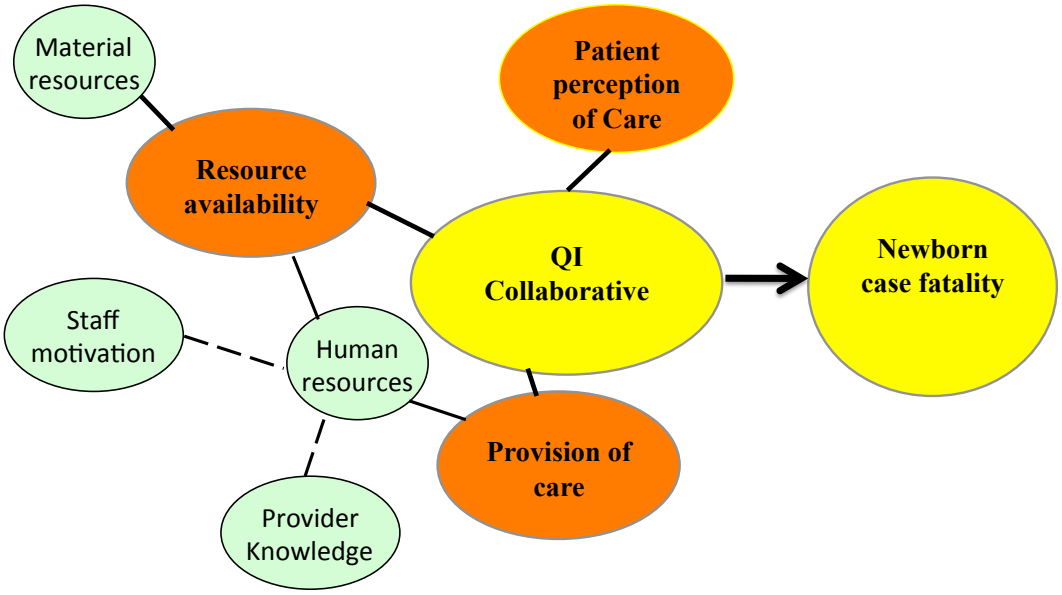

Figure 6: Overview of MaiKhanda process evaluation

The availability of resources includes the availability of physical infrastructure, human resources, and the availability of drugs and supplies. The availability of drugs, supplies and BEmOC and CEmOC signal functions was assessed on a monthly basis using the Health Facility surveillance tool. The availability of human and material resources was determined by the Health Facility Resources Survey (§Appendix-2.3). Measures such as the nurse to population ratio; provision of water, electricity and referral services, were also looked at in this survey.

In addition to the availability of human and material resources, the provision of care is dependent on the level of clinical knowledge and skills of health facility staff and their motivation to work under adverse conditions. Provision of care also includes appropriate management of emergencies and the referral system, and collection and utilization of good data for decision-making. Motivation and morale of staff is analysed using the Staff Psychology survey and clinical knowledge of staff by the provider knowledge survey. The actual physical provision of care to patients, which would require direct observation by clinically skilled personnel, was not assessed. However, change package specific to neonatal clinical care such as emergency newborn care and neonatal asphyxia were assessed using the provider knowledge survey.

Patient's experience of care includes respect, dignity and emotional support provided to the mothers at the point of care as well as addressing equity concerns in service delivery. The Women Friendly Care survey provides an assessment of the patient experience of care

1.8.3 Economic Evaluation

The primary focus of the economic evaluation was to determine the cost-effectiveness of the interventions. Bayesian methods were used to estimate the incremental cost-effectiveness of the community and facility interventions on their own (CI, FI), and together (FICI), compared to current practice in rural Malawi(32). Sensitivity analyses surrounding the costs were undertaken in order to provide estimates of the likely costs of a future scale-up of the interventions.

In summary, evaluation design for the MaiKhanda programme consisted of an impact evaluation using a cRCT design to measure maternal and newborn mortality, a process evaluation analysing the QI intervention and its key components for improving quality of care and an economic evaluation looking at the cost-effectiveness of the intervention.

1.9 My role in MaiKhanda evaluation

The evaluation team included Prof.Anthony Costello and Dr.Timothy Colbourn and myself. Prof.Costello provided oversight for the evaluation while Dr.Colbourn provided technical support to the MaiKhanda team and analysed the study results. I was involved in the initial conceptualization and design of the evaluation of this project (in 2005). From 2007 till the end of the project, I was based in Lilongwe being the technical lead for evaluation. I was mainly involved in:

- Developing the trial protocol and get ethical approval from Malawi and from UCL
- Training and capacity building of the locally recruited research staff in Malawi
- Ensuring data quality for the surveillance systems and the various sub-studies conducted throughout the length of the project.
- Developing study protocols for the various process evaluation sub-studies
- Develop tools for data collection for the process evaluation studies
- Planning and monitoring data collection
- Analysis and reporting on the sub-studies
- Interpretation of the trial results

1.10 Thesis Objective

The MaiKhanda study was the first of its kind in a developing country setting, evaluating the effect of quality improvement intervention on newborn mortality using a RCT design for impact evaluation This is a completed trial with known study results(24). The randomised trial showed no impact of either of the interventions on maternal mortality. The trial showed a decrease in newborn

mortality by 22% [Odds ratio=0.78 (95% CI:0.60,1.01; p=0.057)] in areas with both the facility intervention at the health centre and the community intervention in surrounding villages compared with control areas. The perinatal mortality decreased by 16% [Odds ratio=0.84 (95% CI: 0.72,0.97; p=0.020)] in areas with the community intervention only. In areas with purely QI interventions, the study did not record any significant reductions in maternal, newborn or perinatal mortality(24). Absence of effect could be a result of failure of intervention design or a failure of program implementation or a failure of evaluation approaches. Exploration of these key questions is the cornerstone of my PhD thesis.

The data collection for the study trial was completed in December 2010 and the final evaluation concluded in January 2012. The list of process evaluation sub-studies conducted through the length of the project and described in (Figure 6), was included in the final evaluation report to The Health Foundation (33). However, the report did not go as far as explaining the relation between the individual process evaluation studies and the intervention outcomes. I attempt to explore the relationship between quality improvement intervention and the observed outcomes of the MaiKhanda programme, in my PhD by conducting a secondary analysis of all the available process evaluation studies and other relevant programme documents, using a Theory Based Evaluation (TBE) approach.

Theory based evaluation has been widely used in health promotion programs(34) and is especially helpful in evaluating complex interventions. It maps out the causal chain from inputs to outcomes and impact and tests the underlying assumptions(35). In trying to develop a holistic understanding of the MaiKhanda programme and its (non) effect, I use a theory-based approach to evaluate the quality improvement intervention for newborn health in Malawi. This approach was not included in the initial evaluation design of the MaiKhanda programme. However, it is constructed around the evidence from the process evaluation studies conducted at MaiKhanda as well as relevant programme documents and monitoring reports that were generated

throughout the intervention period. This theory-based evaluation is thus, based on secondary analysis of the data generated from evaluation of the MaiKhanda programme.

I have limited the scope of my PhD to only consider quality improvement interventions for newborn health (i.e the women's group intervention and maternal health have been excluded from my study) for the following reasons:

It was recognised, even at the outset of the intervention, that it was an ambitious target to try and measure maternal mortality. Given the relatively low prevalence (675 per 100,000 live births), it was unlikely that the sample size would be adequate to detect reductions in maternal mortality[iii]. We did not see an effect of either of the interventions on maternal mortality probably as a result of limitations to the evaluation design and sample sizes. On the other hand, both the interventions were sufficiently powered to detect a 30% reduction in newborn mortality. So the changes in newborn mortality are more likely to be influenced by the intervention design and implementation.

Secondly, the evidence base regarding the success of women's group intervention in reducing delays 1 and 2 is fairly well established while evidence regarding the effect of quality improvement on impacting delay 3, less so. A rigorous evaluation of QI interventions, especially in low and middle-income countries is lacking.

Moreover, from a health systems research perspective, it is useful to measure a quality improvement intervention using either maternal or newborn (or both) health outcomes act as a 'probe' to understand how improvement works within the broader health system. My decision to focus on quality improvement of newborn health, for the purposes of my PhD was based on

---

[iii] *Excerpts from MaiKhanda Evaluation Report: The randomised trial showed no impact of either of the interventions on maternal mortality and given the observed number of births and deaths in each arm of the randomised trial the probability (calculated using the properties of the binomial distribution) that 30% reductions in maternal mortality really occurred as a result of the MaiKhanda interventions in the total population the sample sought to represent (intervention areas vs. control areas) was only around 1% for the facility intervention.*

the availability of process evaluation data that I needed for secondary analysis. There was an impetus right from the evaluation design stage to explain the 'black box' of the QI intervention and therefore the process evaluation data that was developed for the QI intervention was quite comprehensive, being based on Hulton's quality of care framework(31). In contrast, data for the community intervention was mainly limited to implementation and monitoring.

The purpose of this thesis is to demonstrate the application of alternate research approaches to evaluate complex interventions in low-income country settings.

The aim of my PhD is to understand and explain how (or why) a quality improvement intervention proclaimed to be successful in other settings failed to demonstrate an impact on newborn outcomes, in Malawi.

The main objectives of this thesis is:

- To understand the program theory behind quality improvement interventions i.e the interplay between a QI intervention and its context which triggers a mechanism (or not) to influence the quality of care and ultimately impact on health outcome for the newborn.
- To analyse the relationship between an intervention's implementation strength and the context and complexity in which the intervention is being delivered.
- To analyse the relationship between program theory and implementation theory

Based on the objectives of the study, the thesis will try to answer the following questions:

- To what extent were the interventions actually implemented in the intervention areas (as compared to the original proposal)?

- How did the intensity and coverage of the QI (Quality Improvement) teams through collaborating learning sessions and mentoring and coaching visits affect the improvement efforts in the facilities?
- What were the contextual factors that influenced the mechanism of the interventions?
- What were the factors and underlying mechanisms that might have rendered the intervention to be (or not to be) effective?

1.11 Structure of this document

Chapter 1, the introductory chapter gives an overview of the intervention design as well as the evaluation design of the MaiKhanda programme. In Chapter 2, I briefly touch upon some of the definition and concepts related to quality of care and quality improvement followed by a review of literature on quality improvement in low and middle-income countries. In this chapter I also discuss the challenges in evaluating quality of care. Chapter 3 describes the study trial and the results of QI intervention on newborn mortality, briefly. Here, I also discuss the limitations of conventional evaluation design in evaluating complex health care interventions. Chapter 4 provides a description of the research strategy. Chapter 5 outlines the evolution of MaiKhanda's Theory of Change, followed by Chapter 6, which provides a synthesis of the available evidence. This synthesis provides a comprehensive framework for analysis. Chapter 7 is an analysis of the evidence using theory-based approaches. In Chapter 8, I provide a detailed discussion of the findings and also discuss some of the methodological challenges in evaluating QI interventions. Chapter 9 is the conclusion chapter where I provide a synopsis of the findings and recommendations for future research.

Chapter 2 Quality Improvement: Definitions and Literature review

The purpose of my literature review is to understand how best to evaluate QI interventions. In order to do this, I begin by developing an understanding of what quality improvement in health care means and establish a clear definition of the intervention. The Chapter begins with a brief history of quality improvement followed by a clarification on definitions and concepts related to quality improvement. The next section provides an introduction to the various approaches to quality improvement followed by a detailed description of the model for improvement adapted by MaiKhanda. Next I identify some of the challenges in defining the outcomes for QI interventions. The primary objective of my literature review is to analyse the effectiveness of QI Collaboratives. I do this by reviewing literature on the effects of QI Collaboratives on health outcomes and identifying key determinants that influence the success or failure of these Collaboratives. Finally, having outlined the challenges and reviewed the literature on QI effectiveness, I scope some of the methods used for evaluating Collaboratives.

2.1 Quality in health care

The concept of quality in healthcare has been in existence for a long period. The importance of observing patients and detailing case notes with accuracy was emphasised in the writings of Hippocrates as early as 400 B.C(36). In the pre-industrial era, standards of medicine were already described in the charter of Royal College of Physicians by 1518(37).

Even before the introduction of “germ” theory, Dr.Ignaz Semmelweis in 1847, demonstrated that handwashing with chlorine could successfully reduce maternal case fatality in his hospital in Vienna(38, 39). Similarly, during the Crimean War of 1854 Florence Nightingale addressed the link between hospital sanitation and high fatality rate(40). She introduced hand washing, sanitizing surgical tools, changing bed linen regularly, keeping wards clean and was able to bring down fatality rate in wounded soldiers from 60% to 1%. However, during this period there was very little emphasis on systems that

promote quality. A systematic approach to quality in health care was adapted from 18th century onwards, largely from the manufacturing industry.

The history of quality in manufacturing began with the advent of the factory system in Great Britain in the 1750s, followed by the industrial revolution in the 1800s. Industrial production meant the integration of numerous processes and co-operation among many different individuals to manufacture one single product. The main mechanism to assure the quality of this product was through product inspection(41) giving rise to the concept of quality control.

Quality control gained prominence and became a critical component of the weapons industry during WW II and was subsequently used in the car manufacturing industry. It began with W.A Shewhart from Bell Laboratories who, in 1930 introduced a scientific approach to quality control(42, 43). Using his technique known as 'statistical quality control', the manufacturing process could differentiate between random variation and non-random causes thereby reducing waste and improving the quality of their product(44, 45).

Meanwhile, in the Japanese Toyota Production System, Deming and colleagues had a very different conception of quality-it was based on methods for improvement of the processes rather than control and conformance. The Japanese system was focused on reducing waste (kaizen/muda) within the manufacturing industry and was influenced by Deming's System of profound knowledge and was considered largely successful(46). This led many industries to adapt this alternate model called quality improvement.

Quality in healthcare was greatly influenced by these quality movements in the other sectors(47, 48). Initially, the focus was more on monitoring individual performance of service providers (similar to control through inspection). More specifically it was on provider performance and did not necessarily touch upon provider-client relationships or organizational characteristics.

Adaptation of quality assurance within the health sector was from from 1960s onwards with Donabedian publishing his seminal paper on framework for evaluation of quality in healthcare(49), providing a broad definition of quality

and recommending that it be measured in three areas as structure, process and outcome(50). Donabedian was a pioneer in the conceptualization and measurement of quality in healthcare. He defined quality as being an evaluative dimension of the elements and interaction of medical processes of care. The review of patient care was mainly through measuring provider performance but then other measurements such as client behaviour, provider-client relationships were also considered important.

In 1951 the Joint Commission on Accreditation of Hospitals was established in the US and there was a paradigm shift in the approach- from *minimum* standards of care to *optimum* standards of care(51). The concept of optimum care redefines quality given the limitations of time and resources, within a given setting. This is done through standards and criterion based audits- these audits lead to decisions within the particular context, and feedback is initiated. Clinical audits gained popularity in the early 1970s followed by a gradual shift to hospital-wide quality assurance programmes, towards the end of the seventies.

A brief description of the terminologies quality control, quality assurance and quality improvement is provided in Box 1

One of the major challenges with quality interventions is the ability to

Quality Control (QC) is the application of statistical techniques to a process in an effort to identify and minimize the random and non-random sources of variation.

Quality Assurance (QA) refers to activities which monitor the quality of a service and may include methods to improve the service. Quality Assurance has three components: Structure, Process and Outcome. Structure includes the physical, human resources and organization concerned with the activity of interest. Process includes the acts of delivering of care. Outcome includes the effects of care.

Quality Improvement (QI) is the effort to improve the level of performance of a key process. It involves measuring the level of current performance, finding ways to improve that performance and implementing new and better methods.

Box 1: Brief description of terminologies related to quality in healthcare

associate quality interventions with outcomes(52). Moreover, since the outcome measure (such as quality of life) and process measures are complex and difficult to define, the emphasis is usually placed on evaluation of an element such as provider performance(53). Such a narrow model of quality assurance is usually met with resistance from heath workers for fear of regulatory agencies, litigation and external review(54).

Quality improvement on the other hand complements quality assurance by adopting a coordinated and integrated approach to improving processes that influence patient outcomes. Here employees have a central role to play and the focus is on improving work processes rather than supervising individuals(55).

In 1988, the Joint Commission adapted the technique of continuous quality improvement and advocated for a multi-disciplinary approach that focuses on improving the performance of an entire group and shifting the entire production curve upwards(54) rather than identify isolated poor performers.

While the origins of quality improvement approaches in healthcare can be traced back mainly to the US, quality is an inherent part of the healthcare system within the NHS, in the UK as well. Quality movement in the UK was aimed at accountability and providing quality care and at the same time optimizing resources(56). In an effort to improve quality, the NHS now has a audit of provider performance(57), which as mentioned earlier, offers limited understanding of the quality of health services provided.

Quality in healthcare in Africa and other low-income countries is a fairly recent concept, influenced largely by the quality movement in healthcare in the US and UK. In Africa, most of the quality improvement projects have been established through USAID's flagship Health Care Improvement project(18, 58, 59). The QI movement in UK offers much to learn for the adoption of QI in most of Anglophone Africa, which have publicly funded health systems similar to NHS in UK.

2.2 Quality Improvement: definitions and concepts

Lee & Jones in their 1933 classical work -The Fundamentals of Good Medical Care defined quality as "… the kind of medicine practiced and taught by the recognised leaders of the medical profession at a given time or period of social, cultural and professional development in a community or population group"(49, 60).

Donabedian expanded it further to include "the application of medical science and technology in a manner that maximizes its benefit to health without correspondingly increasing the risk"(61, 62). In evaluating healthcare improvement, Donabedian defined quality in terms of structure, process and outcomes. This involves evaluating health systems capacities and the interaction between clinicians and patients as well as evidence about

changes in patients' health status(63). While all 3 dimensions are important for measuring quality, most of the literature on quality is focused on process of care, more specifically, assessing appropriateness of care (mainly through provider performance) and adherence to standards.

The IOM (Institute of Medicine) 2001 report identified the “quality chasm”-the big gap between care that people should ideally receive and what they actually receive. Given the complex notion of quality, the report defined quality in terms of its six dimensions(63):

Patient centeredness*,* i.e. providing care that is respectful and responsive to individual patient preferences, needs and values.

Equity: Health care systems have to improve the health status of a population and at the same time reduce the disparities between them

Effectiveness-providing services according to current best scientific knowledge

Efficiency such that resources are used in a manner to get best value for money. This could be done through reducing ‘waste’ and reducing production (or administration) costs.

Timeliness i.e avoiding harmful delays

Safety avoiding injuries for patients from the care that is intended to help them.

The WHO definition of health care is broader in scope as it encompasses both the health care as well as the health system in which it is operating. It adopts a system-wide perspective in its approach to quality and the focus is on organizing the care delivery system and not just on medical knowledge and resource availability(64). Using a system perspective, quality can also be defined in terms of roles and responsibilities on individuals, teams as well as organizations.

As is clear from these various definitions, quality is multi-dimensional. In simple terms, Quality Improvement is about improving the performance level within these various dimensions of quality.

The definition of quality improvement specifically within MNH (maternal and newborn health) covers specific dimensions around reproductive health rights, effective and timely access to care, including provider needs(31, 62)

2.2.1 Core principles of QI:

Langley et al and others have identified the core principles underlying a quality improvement intervention(46, 48, 65). Key principles include:

- Placing patients at the centre and involving them in the co-design.
- Understanding work processes as components of a wider system and re-designing accordingly.
- Improving the reliability of the system and clinical processes
- Understanding variation and measuring the processes.
- Using data for measuring improvement.
- Recognizing and valuing the expertise of people in the frontline. They provide a broader definition of improvement science, which includes commitment to practical learning, generating local wisdom, contributing to clear and explicit theories of how change happens.
- Focussing on the design, deployment and assessment of complex multi-faceted interventions(66, 67).

2.3 Quality Improvement Framework

Quality is an integral part of patient care but reference to the terms quality control and quality assurance implies an external agent evaluating the process and providing feedback. Quality improvement on the other hand is "*the continuous and relentless effort by everyone to improve patient outcomes, improve system performance and capacity building*"(68).

The quality improvement framework that was used in the MaiKhanda programme was based on IHI's Model for Improvement and is presented in detail below and diagrammatically represented in Figure 9. It consists of the following components, each of which is described in the sections below:

- Deming's system of profound knowledge

- Various approaches to quality improvement (with an emphasis on Rapid Cycle Change which was the key approach used in the MaiKhanda programme).
- Model for Improvement
- QI Collaboratives

2.3.1 Deming's system of profound knowledge

For improvement to occur, it involves change and this change is influenced not just by the available scientific evidence but also the degree of belief that people have in that evidence. This intersection between belief and best available scientific evidence is at the heart of quality improvement(69). The degree of belief stems from a system of profound knowledge as described by Deming. He describes profound knowledge as 4 independent processes, all related to each other (Figure 7):

- Appreciation of the systems
-

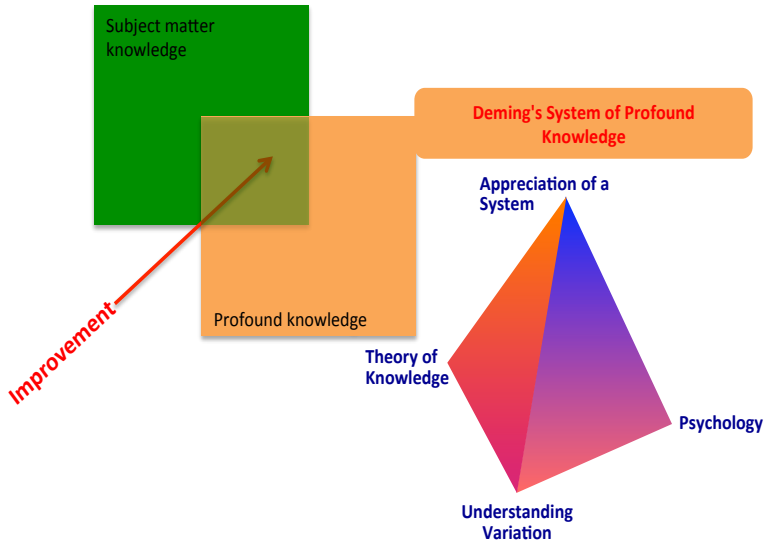

Figure 7: The science of improvement

A system is an interdependent group of items, people or processes working together towards a common purpose. The success of an organisation depends not only on the performance but also on integration of the individual

parts. Multiple measures are therefore required to understand the impact of the changes on the various individual parts as well as whole system changes. Systems are not made of simple linear cause-effect relationship and changes in one part of the system may produce unpredicted effect in another part or at a later time. Unintended consequences are a common experience and leads to misinterpretation of results.

- Understanding variation
-

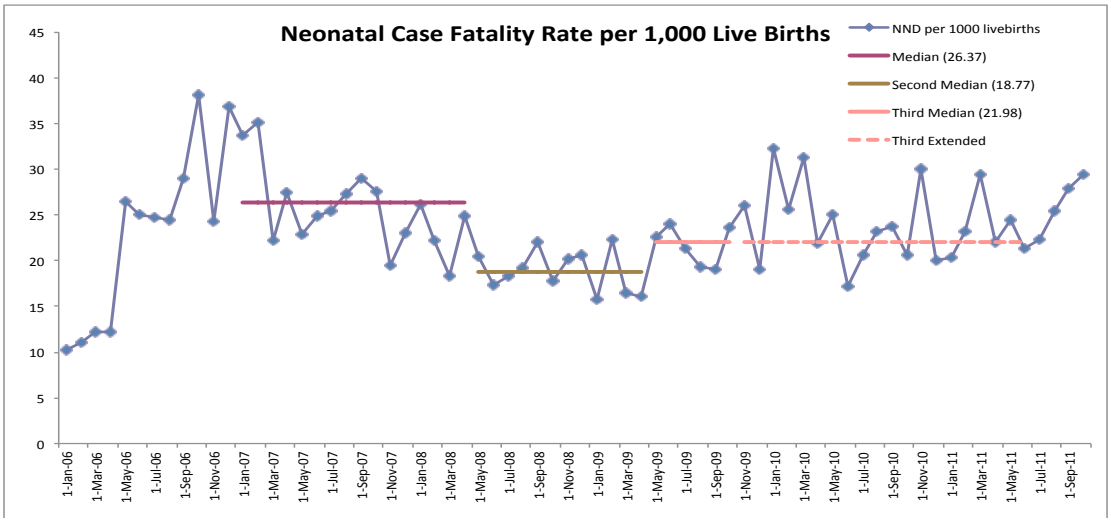

Figure 8: Example of a run chart showing neonatal case fatality rate in MaiKhanda's 9 CEmOC facilities

Everything that is observed or measured varies. This can be understood by plotting data over time. In 1920, Walter Sehwhart pioneered the theory around variation(43) and introduced SPC (Statistical Process Control) mainly to improve manufacturing. SPC has been used in healthcare since 1960s, initially in laboratory settings and then eventually for direct patient care application(70). Variation can be differentiated as common cause variation

where patterns in the data are predictable and special cause-these are variations which lead to fundamental changes within the system(71).

Processes that affect common cause within the system are called stable processes, while those processes whose outcomes are affected by both common and special cause are called unstable processes. It is common to mistake a common cause variation to special cause and tweak the system or a special cause variation as common cause and miss the opportunity to make a change. The commonly used tools for understanding variation are run charts (Figure 8) and control charts(72, 73) developed by Shewhart in the 1920s to improve industrial manufacturing(43)

The control chart consists of 2 parts: a series of measurements plotted in time and the control chart 'template' which consists of a central line (usually the mean or median) with upper and lower control limits. When data falls outside the control limits it is interpreted as a 'special cause' variation that is variation caused by some inherent changes in the process. These changes in the process can be the result of a deliberate intervention or a natural event over which we have little control.

- Building knowledge

From an improvement perspective, a change is a prediction. So the greater the understanding and knowledge one has about the functionality of a system, the better the prediction and greater the likelihood of change resulting in improvement. Comparing predictions to results is a key source of learning. Rational prediction requires theory. Theories are built using current knowledge and then designing tests to validate these theories. These tests are done using the PDSA (Plan-Do-Study-Act) cycle(46). In improvement, more commonly used is the repeated deductive-inductive cycle of Plan-Do (deductive approach) Study-Act (Inductive approach). The tests help identify gaps to the prediction, and theory is updated accordingly. Action is then taken on the new learning.

- Theory of psychology (the human side of change)

Knowledge of the human side of change helps to understand how people interact with each other and with the system. It helps understand the motivation of people and their behaviour(71).

In summary, improvement occurs when there is overlap between subject matter knowledge and profound knowledge(46). From an operational perspective, 'improvement science' defines this as the use of scientific principles and methods to address the practical challenges of delivering better health care for patients and populations(74). Quality Improvement aims to narrow the gap between what we know from research i.e. the practice of evidence based medicine and what is done in clinical practice by using quality improvement methods. In essence, the focus of evidence based medicine (EBM) is doing the right things and that of quality improvement (QI) is about doing things right (in the right way)(75). Evidence based medicine through systematic reviews, practice guidelines and computer assisted clinical decision-making tries to improve the quality of evidence on clinical care that patients should receive(63). But that in itself will not address the gap in existing practice. Implementing EBM to improve quality of care requires understanding and adaptation to the context in which the guidelines are being introduced.

In order to put improvement in practice, a simple 'Will-Ideas-Execution' framework has been suggested that lays the foundation for an improvement to be initiated. This includes the *Will* to do what it takes to change a system, *Ideas* on which to base the design of the new system and *Execution* of those ideas(76). There are various approaches to putting improvement into practice. They are briefly mentioned in the next section.

2.3.2 Approaches to quality improvement

Powell et al(77) (NHS QI Scotland 2009) provide a systematic narrative review of the various QI models in health care(77). They identified 5 models of quality improvement:

- Total quality management

- Business Process Re-engineering
- Lean thinking
- Six Sigma
- Rapid cycle change

2.3.2.1 Total Quality Management (TQM)

This is also known as Continuous Quality Improvement (CQI). It was developed in Japan in 1950s and used in health care since 1990s. The emphasis is on quality improvement as an on-going activity. It is data driven and has an 'empowered' mutli-disciplinary team at its core. TQM considers quality improvement as a normal and integrated on-going activity within an organization.

2.3.2.2 Business Process Re-engineering (BPR)

The emphasis is on the importance of examining and redesigning the processes, e.g. re-designing care pathways in the NHS. It is usually driven from the top by visionary leadership. Contrary to other improvement approaches BPR does not focus on incremental changes-it adopts an “all or none approach”. BRP in health care has evolved in different ways and has in most instances been implemented partially.

2.3.2.3 Lean thinking

This was a concept developed by Toyota in the 1950s. The principle behind this model is streamlining processes with minimal wasted time, effort and costs. It consists of value stream mapping that is analysing current processes to generate ideas for process re-design in order to remove any unnecessary steps in a process. It has been applied in health care settings with some success in reducing waste(78). The approach appears to be particularly useful in streamlining processes in support departments rather than mainstream clinical services. Wholesale application has not been demonstrated in a health care setting.

2.3.2.4 Six Sigma

One of the recent additions to quality improvement, which has been in use since the 1980s. It follows a structured approach known as DMAIC (Define-Measure-Analyse-Improve-Control) to eliminate waste and reduce variation in the processes so as to improve outcomes. It is dependent on extensive use of statistical tools such as SPC (Statistical Process Control). It has only been applied to a limited extent in healthcare primarily because the approach requires statistical expertise alongside reliable local data collection. It is more commonly used in combination with lean thinking.

2.3.2.5 Rapid cycle change

This is based on the model developed by Nolan and Langley(46) for improving organizational performance in multinational companies as well as different industries such as healthcare and public agencies. The rapid cycle change consists of repeated short-cycle small-scale tests of change linked to reflection, known as Plan-Do-Study-Act cycles. This includes conducting repeated cycles of quality problem diagnosis (PLAN), development and implementation of small-scale improvement efforts (DO), assessment of effects (STUDY), and refinement and expansion of effective actions (ACT) until desired outcomes are achieved. PLANning involves outlining the objectives and determining the questions and predictions as well as planning for data collection. DO involves actual execution of the plan and documenting the observations. STUDY involves the analysis of data and comparison to the predictions. Data is summarized and discussed by the QI team so as to ACT on the changes that are to be made or plan for the next cycle.

The rapid cycle approach was implemented in Malawi as part of the MaiKhanda project(24).

Improvement interventions usually use a hybrid of various improvement techniques. One such hybrid is the QI Collaborative approach which was first introduced by IHI through its 'Model for Improvement' in 1996 (79, 80).

2.3.3 IHI's Model for Improvement

The model for improvement comprises of 3 basic questions that QI teams have to focus on, in order to develop their PDSA(44) cycle. The questions are:

- What do we want to achieve? This includes defining the aim of the project. The aim is to be succinct, time-bound and measureable.
- How will we know that the change is an improvement? All improvement involves change, but not all change is an improvement. Improvement is usually measured using tools such as run charts.
- What change can we make that will result in an improvement? This includes ideas to bring about change.

Once the teams have answered these preliminary questions, they get working on the PDSA cycles.

In order to help the QI teams answer these preliminary questions and develop their PDSA cycles, they are supported by improvement advisors, who provide the necessary mentoring and coaching to fulfill their objectives. The teams also receive encouragement and motivations through inter-organizational networking. All this is achieved through the QI Collaborative. The QI Collaborative consist of 2 distinct parts: the Collaborative 'Learning Sessions' where organizations get a chance to interact with each other and learn from each other's improvement ideas and an 'Action Period' where facility QI teams are visited by improvement advisors and receive onsite coaching and mentoring.

The length of Collaboratives varies over time depending on scope of the intervention, and the collaborative is characterized by a continuum of learning sessions and action periods. Thus the Collaboratives are a combination of PDSA and inter-organizational networking

2.3.4 The QI Collaborative

The "QI Collaborative" (QIC) is also known as the "Break-Through Collaborative Series"(77, 79-81).

The Collaborative consists of groups of participating health care delivery organizations. Each organization is represented by a 3 or 4 person team who are members of the QI team in their respective health facility. The teams are supported by a faculty of QI experts. The objective of these Collaboratives is to study a specific health care issue, design and implement specific solutions, evaluate and refine these solutions, and disseminate findings to other organizations.

Method: The Collaboratives consist of a 'learning session' which is a series of 2-3 day meetings over a period of several months. During the meetings, the team members with support from the faculty experts, learn improvement techniques, exchange insights and advice, and generate enthusiasm and a shared sense of commitment to improvement goals and outcomes.

Followed by the Collaborative session is the "action period" when teams return to their organizations between learning sessions to apply their new knowledge and ideas in a Plan–Do–Study–Act cycle(82) using the improvement model framework(46).

.

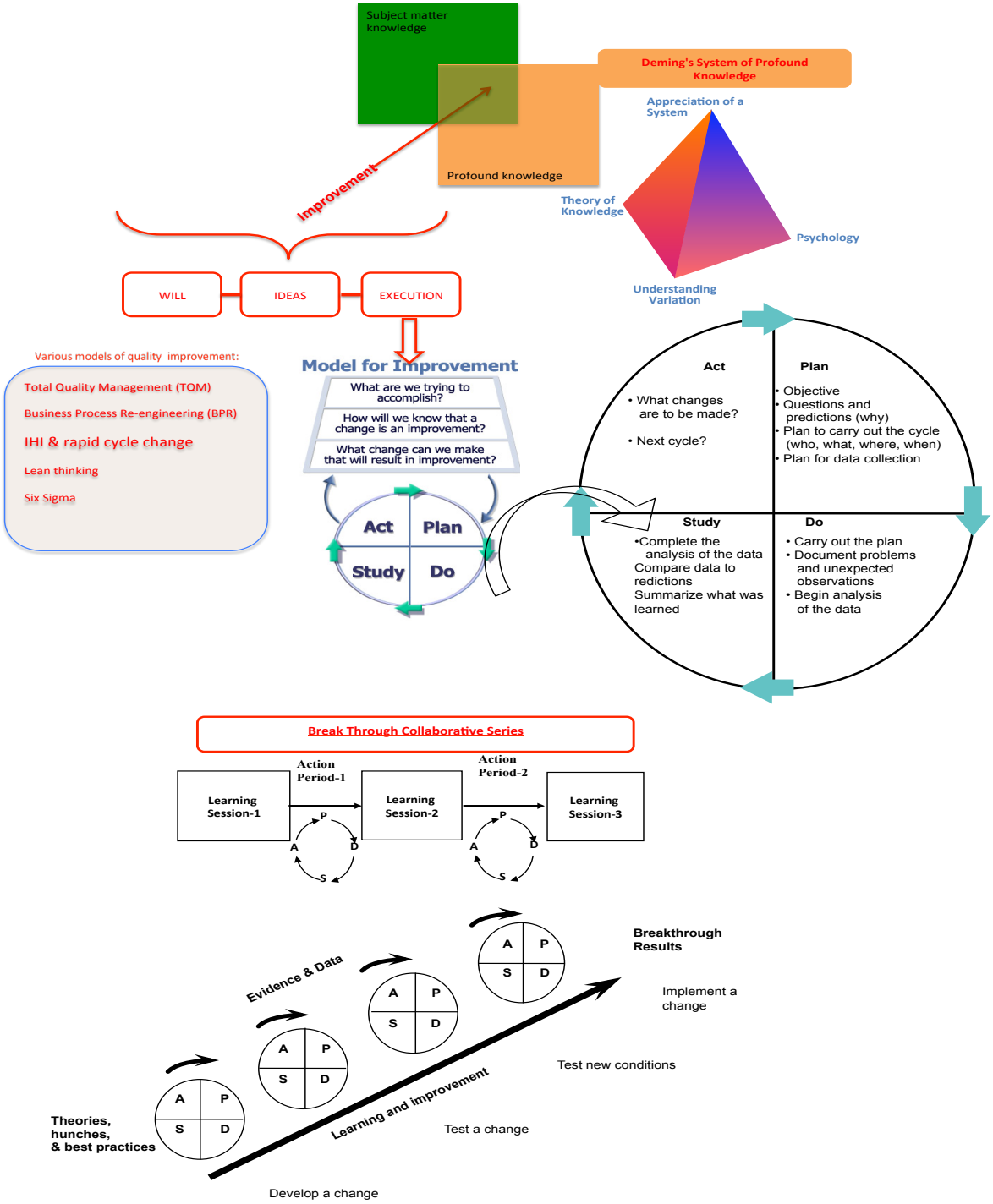

Figure 9: Quality Improvement Framework

Figure 9 summarizes the various concepts described above. To summarize, QI Collaborative is an organised multi-faceted approach to QI that involves 5 essential features:

1. Specified topic: this is usually a clinical or administrative subject that addresses gaps between the best and current practice. Improvement

experts provide with the evidence and change ideas, which saves time on the initial investigation work(83).

2. Ideas and support for improvement: clinical and QI experts identify, consolidate, clarify and share scientific knowledge, best practice and improvement knowledge
3. Critical mass of multidisciplinary team from multiple sites
4. Model for improvement consisting of clear and measurable target, data collection, small tests of change, learning by doing
5. Collaborative process involving a series of structured activities within a given time frame, exchange of ideas and sharing of experience between participating teams.

2.3.4.1 Strengths of Collaboratives

One of the expected outcomes for using the collaborative approach is for the team to take 'ownership' of the project or event, a necessary pre-requisite for successful organization change(84).

These methods of testing small tests of change are designed within the local context and are therefore, arguably, sustainable in the long run. It draws on the ideas and ingenuity of local staff. Because of its localized nature, the collaborative enables low risk testing of changes in the clinical setting, it can be used relatively informally, and can be used to address issues at different levels within the health system.

The Collaboratives provide a sharing and learning environment where teams are able to share successes and failures and learn from their peers at the same time(85). Collaboratives adhere to most of the core principles of a QI initiative (§2.2.1).

2.3.4.2 Weaknesses of Collaboratives

Findings from the NHS Orthopaedic Services Collaborative provides insight into a weak collaborative(86). Of the 37 Trusts involved in the Collaborative, 7 Trusts withdrew, there was a high turnover of project managers, there was partial implementation of PDSAs, limited evidence of networking between

groups, lack of resources allocated to the process, insufficient attention to developing receptive context locally and a general lack of interest from many participants. In addition, conflict between teams and organizational strategic objectives as well as conflict between different projects within the same organization have been identified as potential weaknesses(87, 88). Single changes could displace problems on to other parts of the system. There is a risk of partial use of PDSA methods providing premature solutions(70). For Collaboratives to be successful there has to be the will and skill for data analysis, accountability for results and sustaining efforts(89) .Obtaining robust data is a challenge. There is limited evidence from peer reviewed literature in terms of change in outcome(18) and PDSAs could be ineffective when dealing with intractable systemic or bureaucratic problems(90).

2.4 Geographic scope of QI Collaboratives

Some of the earliest known Collaboratives were initiated in North America and include the Northern New England Cardiovascular Disease Study Group (1986) and the Vermont Oxford Network (1988)(91, 92). The first Breakthrough Series Collaborative was initiated by IHI in 1996(79). Since then Collaboratives have been used in a wide range of health care setting in the US(93) and elsewhere.

Collaboratives[iv] are widely used in the NHS. The UK cancer collaborative was one of the first to be established in 2002. Others include the National Primary Care Collaborative, and the Scottish Primary Care Collaborative programme(77).

QIC are widely used in Canada, Australia, Sweden, Netherlands, Norway(94).

Quality improvement has been a recent introduction in the developing world. The majority of the QI projects in low and middle income countries have been implemented through the USAID funded health care improvement project(18,

[iv] *This refers to Breakthrough Collaborative Series in the remainder of the document.*

59). A literature review on quality improvement in Africa done by Liu and colleagues yielded 10 studies(58).

2.5 Literature review: Quality Improvement Collaboratives

2.5.1 Defining what needs to be reviewed

The primary objective of any quality improvement intervention is to improve patient outcomes through improving quality across the inter-dependent levels of the health system (23, 95). This includes:

- The individual level which includes the patients but also the various health care providers and other individuals involved in the direct or indirect provision of care.
- The microsystem: This includes the clinical system(s) at the health facility for eg; maternity and labour ward, laboratory systems, data management systems. Most of the QI interventions act at this level of the health system.
- The macrosystem: This consists of the broader health care organization, the referral systems between different levels of the health system, the Ministry of Health (MoH), the District Health Management Team (DHMT).
- The environment: This is the broader context in which the micro and macrosystems are functioning and includes external stakeholders, the policy environment, the global context for (in the case of MaiKhanda) maternal and newborn health.

The purpose of evaluation then is in understanding and ascertaining the relationship between the process changes that happens across the different levels of the health care system and variation in outcomes(96). The question of what to measure in evaluating quality improvement is important since the parameters chosen have an influence on the interpretation of results. Figure 10 represents an outline of an evaluation framework for QI interventions which is a modification to Donabedian's framework of structure-process-output for evaluating quality interventions(49). I begin this section by discussing the

measurement of outcomes (Figure 10), which focuses on the clinical (or health) outcome of the individual patient. There are usually includes 3 kinds of outcome measures: biological (such as mortality and morbidity measures), behavioural and psychological. Mortality is the commonly used outcome measure for most interventions as it is a very objectively verifiable indicator. Irrespective of the kind of outcome measure chosen, there is an assumption that improvements in quality of health care services will lead to measurable changes in health outcomes.

Since these individual outcomes are complex to measure, a more commonly used approach is to evaluate quality of care itself. However, quality of care cannot be measured as an outcome per se since there is no standard acceptable definition of quality of care. Evaluators usually use one or more of the dimensions of quality of care namely patient centeredness, timeliness, equity, efficiency, effectiveness and safety (IoM 2001) (63) to define and measure quality of care. At an individual level, evaluating quality of care includes evaluating the provider behaviour as well as client behaviour and the complex network of formal and informal relationships between the client and the provider (50).

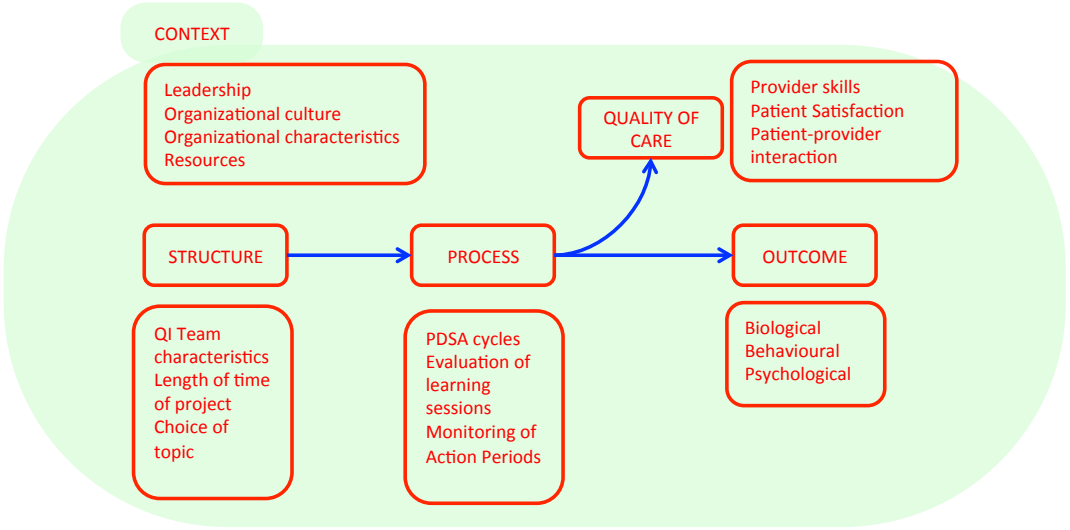

Figure 10: QI Evaluation Framework

In the earlier on period of measuring quality of care, evaluating provider performance was used more frequently and commonly in advanced health care systems to increase transparency and accountability, while client experience (patient satisfaction) is now increasingly being considered in the evaluation of quality of care (53). It has also gained wide acceptance given the changing role and expectations from care providers. Health care providers are expected by their organizations to accept responsibility not just for the clinical quality of care but for other dimensions of health care system such as efficiency and equity. Similarly, technology is becoming an integral part of the care process as well as the patient provider relationship(97) and contributing to patients knowledge and expectations in terms of good quality care. These dimensions also need to be taken into consideration in evaluating quality of care.

The challenges of evaluating individual outcomes or even quality of care as its proxy is not without limitations(49). Since it is difficult to attribute changes in health outcomes to quality of care exclusively, process evaluation can be helpful in explaining the intervention factors that may have influenced the outcomes. Improvement by its very definition involves change. Another important dimension of QI interventions is adaptability(98). Both these concepts are difficult to absorb within existing impact evaluation methods that solely look at hard outcomes such as mortality(99). Process evaluation helps to shed light on if the intervention actually did not show an effect or if the absence of an effect was due to the design or implementation of the intervention being sub-optimal(67). Quality improvement is a continuous and dynamic process constantly evolving over time and therefore evaluation measuring 'processes of care' parameters are more easily measurable than outcome parameters(100). This includes measuring improvements in the change package, monitoring PDSAs (Plan-Do-Study-Act). It can possibly be measured in terms of the QI implementation process using routine intervention monitoring data. Although the use of routine QI data for evaluation purposes requires further scrutiny, it is still an indispensable element of the intervention itself because it provides feedback concerning the

performance of the system. While process evaluation might be able to shed light on the issue of attribution of the intervention to the outcome of interest, reporting on process measures alone, is of limited value as it does not tell us about the effectiveness of the intervention.

Applying a biomedical research analogy of a dose-response relationship, process evaluation can be considered akin to measurement of the 'dose' of an intervention, within public health interventions (also referred to as the programme implementation strength). Measurement of implementation strength of processes consists of two components namely implementation quality (also referred to as fidelity or integrity) and quantity (also referred to as dosage). Fidelity is defined as *the extent to which delivery of an intervention adheres to the protocol, guidelines or programme model originally developed* while the dose, duration, intensity and specificity (often referred to as 'dosage') of a QI Collaborative includes monitoring of Collaborative workshops and action periods, change packages and PDSA cycles(98).

The evidence regarding evaluating implementation strength of complex multi-dimensional interventions is limited, especially from low and middle income countries(98). Furthermore, higher implementation strength need not necessarily produce better outcomes(101). This poses the first of many challenges for evaluation-that is establishing a causal link between improvements in quality of care being provided and direct measurable changes in patient outcomes. This linkage is difficult to establish and especially with complex social interventions because of the low signal to noise ratio - outcomes are influenced by factors other than quality of care(102) such as the nature and magnitude of need, provider-client relationships and the capabilities of medical science(49) as well as the overall context (characterized by the interdependent layers of the health system) in which the interventions are being implemented. It is difficult to link health outcomes with quality unless all these other factors are also considered.

Evaluation is not just focused on summative health outcomes assessment but also the formative i.e. to assess the extent to which an implementation is

effective in a given context(103). Thus evaluation of a QI initiative could have measures looking at the effect on health outcomes or quality of care or the processes of care, but equally important for evaluation is to understand the embeddedness of the intervention within the health system. This requires evaluation of the broader context in which the intervention is being implemented. Thus evaluation of the context is another important aspect of evaluation of QI Collaboratives and focuses on other levels of the health system which will ultimately impact on individual patient outcomes. An essential feature of most interventions is that the events and processes occur within a particular setting. The characteristics of these settings includes formal and informal organization as well as social, economic and cultural factors, which profoundly influence all the elements of the intervention. Furthermore, it has been suggested that these elements interact with each other producing a multiplier (rather than additive) effect(104). Evaluation of quality and performance in health care need to look at these 'cultural transformations' alongside individual structural and process changes(105).

More than 50 contextual factors have been identified by Kaplan et al that could potentially influence QI interventions(106). Some factors were more frequently examined, others less so. Of the commonly examined factors, organizational characteristics, size, organizational culture, organizational maturity (years involved in QI) were generally shown to influence QI success(106). Among the less examined factors influencing QI success include leadership for quality improvement, clinical integration across department, customer focus, microsystem motivation to change, resources and QI team leadership. According to Lombarts et al, the maturity of an organization's experience with QI (measured in terms of an organization's formal involvement in QI) showed a significant positive relationship with QI success(107). Similarly presence of an organizational culture supportive of QI, which was measured using tools such as the competing values framework, showed a positive association with QI success(108). In summary, organizational leadership, culture, QI maturity, and data infrastructure have the strongest evidence base for successful QI interventions in existing

literature(106). Factors such as motivation, leadership and culture operate within and across the different levels of the health system.

Contextual factors related to the clinical microsystem and the QI team are directly related to the success of QI while organizational factors and external environment are believed to influence QI success indirectly(109). But, it is important to acknowledge that it is not just the influence of individual elements of the intervention and the health systems, but also the interaction between them that influences quality of care(65, 109).

In summary, the evaluation of QI Collaboratives needs to take into consideration the structure and processes of care as well as quality of care and their influence on health outcomes. Evaluation also needs to consider the external and internal context in which care is being provided and the inter-relationship between them that influences outcome. Each of these parameters evaluated separately and in combination offers great insight into intervention implementation and its sustainability and replicability.

At this point, it is important to bear in mind that the quality of care is the end result of a QI intervention whereas the QI Collaborative is a technique to improve quality of care and health outcomes. QI *Collaboratives* then are a means to achieve quality of care in the facilities using a combination of inter-organizational networks and learning cycle(86). It is an approach to learning and implementing quality of care using a combination of subject matter knowledge and profound knowledge. Evaluation of QI is many a time focussed on the implementation components such as the QI Collaboratives, or PDSA cycles. These are essential but by themselves are unable to provide insight into the success (or failure) of the intervention.

Given this understanding of the challenges and complexities of evaluating QI Collaboratives, the next section reviews the available literature to outline design and methods that might be best suitable for evaluating QI Collaboratives.

2.5.2 Literature review process

I focus on Institute for Healthcare Improvement (IHI) Breakthrough Collaborative Series(110), as this was the main approach used for the QI interventions in the MaiKhanda programme in Malawi.

I searched for literature on the OvidSP Medline search engine in the following databases-Medline, Embase, PsycInfo and HMIC (Health Management Information Consortium) using both MeSH and non-MeSH terms. The terms included:

Quality improvement [OR] quality assurance [OR] quality of healthcare [OR] continuous quality improvement [AND] collaborat* [or] “breakthrough series”

Separate searches were also done in PubMed and Cochrane database. The search strategy was restricted to English language and there was no restriction on the time period. This search strategy yielded 766 articles. Given the limited number of articles on the topic I did not restrict the literature search to a particular clinical outcome such as maternal or newborn mortality as this would have generated very few hits to further explore the determinants of successful QI initiatives. Moreover, as discussed above, the issue of outcomes in QI initiatives is a widely debated topic in improvement research.

After taking out duplicates (32) there were 734 articles left (Figure 11). Based on the search strategy, I only selected those articles that contained either “quality improvement” or “Collaboratives” in the title. This excluded 623 articles, leaving 111 articles to be shortlisted through their abstract. Forty abstracts that did not meet the inclusion criteria (e.g. not Quality Improvement intervention or QI Collaborative, no measurement of outcome) were excluded. Articles which discussed the design and implementation of the Collaborative methodology and reported on the evaluation of the QI Collaborative were included in this review. Thus 71 full text articles were selected, of which 2 were excluded because their Collaborative model was different from the Breakthrough Series Collaborative considered in this review. I conducted a manual search of references from the studies included. The search strategy

was further extended to include other prominent websites such as the Health Foundation (www.health.org.uk) and Institute for Healthcare Improvement (www.ihi.org). A total of 69 articles was selected through the search strategy and in addition 1 article was identified through the IHI website and 8 were references from the selected articles. After reading through the final list of 78 articles, I identified 15 articles that were critical to this review.

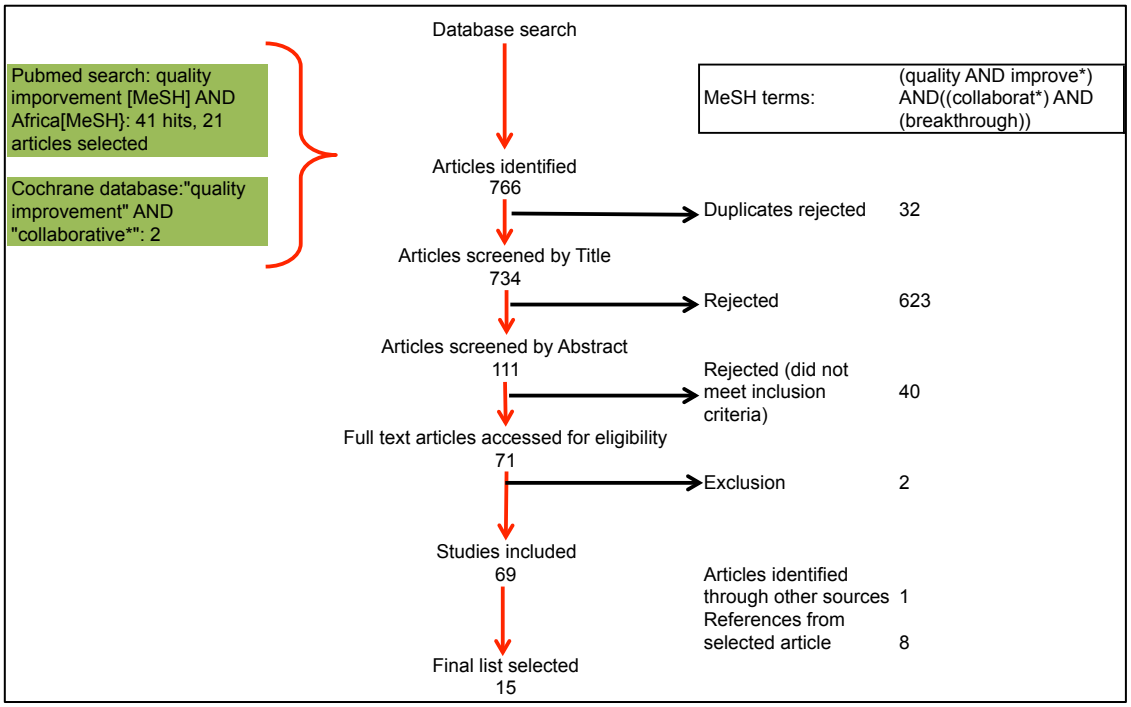

Figure 11: Literature review flowchart

2.5.3 Effectiveness of QI Collaboratives

The model for quality improvement developed by the US Institute for Healthcare Improvement has been widely used as a model for improvement Collaboratives especially in the USA, UK, Europe, Australia and also recently in South Africa[v]. However, there have been few evaluations using rigorous methods (such as RCTs) to measure effects of quality improvement initiatives published in the academic literature, especially for low and middle-income countries.

[v] *http://www.ihi.org/Pages/default.aspx*

Figure 12 presents a summary of the literature covered in this review, organized according to the QI evaluation framework discussed above. A review by Sifrim and colleagues identified 76 articles from low- and middle-income countries that reported on quality improvement interventions, though not specifically on QI Collaboratives(111). These were mainly observational studies. Their characteristics included description of the problem statement, the improvement processes and tools, a description of the interventions itself and their evaluation methods(111). 60 of the 76 studies (79%) did not have a concurrent comparison or control group in their evaluation design. This literature is limited in scope since it does not analyse the effectiveness of the interventions (but does provide some information about evaluation methods).

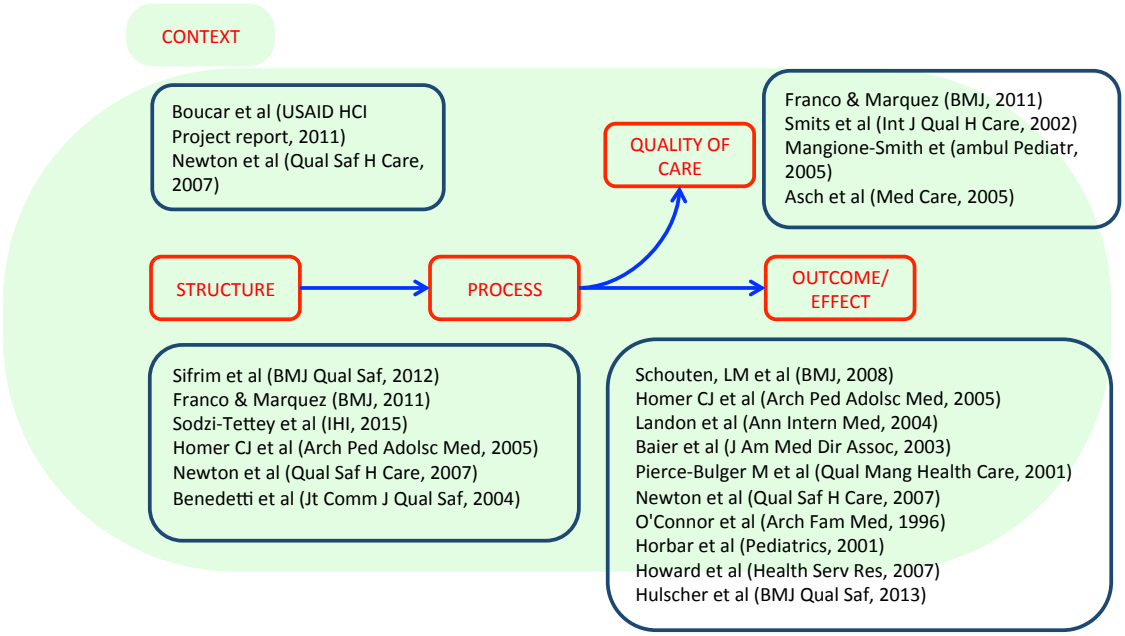

Figure 12: Summary of reviewed articles

Franco and Marquez detail time-series evidence in support of quality improvement breakthrough series Collaboratives leading to improvement in adherence to essential standards of care and some health outcomes in 12 low and middle-income countries(18). They reviewed 54 Collaboratives in 14 low and middle-income countries and found 27 Collaboratives in 12 countries with

useable data, and found evidence supporting the achievement of compliance with standards, and of performance outcomes for 87% of 135 indicators measured in time-series charts. This review was also limited by a lack of comparison groups; it was also limited by the outcomes being self-reported and only of process measures.

Table 3: Summary of effectiveness of QIC from 7 controlled studies

Controlled Studies (n=7)	Breakthrough series	Results	Breakthrough series with chronic care model	Results
RCTs			Homer et al (112)	No effect
Before and After Studies	Landon et al (113)	No effect	Benedetti et al(114)	Mixed effect
	Baier et al (115)	Positive effect	Mangione-Smith R et al (116)	Mixed effect
			Asch SM et al (117)	Mixed effect
Interrupted Time Series	Pierce-Bulger et al (118)	Positive effect		

A systematic review by Schouten(20) and colleagues on the evidence for quality improvement Collaboratives identified 57 studies (from a total of 72 studies) that were based on the Breakthrough Collaborative series . Of these, 50 had an uncontrolled study design and mainly used self-reported measures of improvement. Of the 7 controlled studies (Table 3), one was a randomized controlled trial evaluating the breakthrough series integrated with the chronic care model. The trial did not show an effect on any of the key processes or intermediate outcomes of care for children with asthma(112). The studies had all different measures of outcomes. For instance, Landon et al looking at HIV infection analysed virological outcomes and quality of care provided and showed no significant effect(113). Baier et al showed a significant decrease in the prevalence of patients with pain in 21 nursing homes(115). The interrupted time series by Pierce-Bulger et al showed significant improvement in neonatal deaths(118). While still others such as Bendetti et al looked at

improvement in process measures (eg: annual rates of eye and feet examination for diabetes), while Mangione-Smith et al reported on quality of care such as patient education and self management. Asch et al reported significant improvements in counselling and education in the care of patients with chronic health failure.

The USAID sponsored Health Care Improvement (HCI) Project has nearly 15 projects across Africa working mainly on HIV and related issues(59). The HCI project in Niger(22) focused on institutionalization of quality improvement for essential obstetric and neonatal care and introduced change packages across different levels of the healthcare system. The assessment looked mainly at performance measures but did not report on primary outcomes. A phased, rapid national scale-up quality improvement intervention focused on under-five children in Ghana, tested process changes using QI techniques, along the continuum of care from pregnancy to age 5 in both inpatient and outpatient settings but reported on effectively scaling up MNCH interventions to national level but, did not test effectiveness of the interventions themselves(19, 119).
Newton and colleagues published a non-systematic review of Collaborative methodology and its use in heart failure research. The review identified 43 articles. It concluded that organization and leadership were important aspects of Collaboratives, that Collaboratives improved outcomes of patients (with heart failure and chronic cardiovascular diseases) and that they facilitate sustainability(90). Most of the studies were observational studies, except three which were case control studies. The studies mainly reported on the determinants of a Collaborative (§2.5.4). Some of the studies also reported on changes in process and outcome measures, including costs. Only one case control study reported statistically significant improvement in outcomes (i.e reduction in mean HbA_{1c} levels) (92, 120)

Hulscher et al (2009) further explore the effectiveness of Quality Improvement Collaboratives reviewed by Schouten, by analysing the determinants of their success or failures(121). They were able to identify 13 articles from 10 controlled studies in their review. Eight of the studies were based on the

Break Through Series model developed by IHI. Four of these also incorporated the Chronic Care model (CCM) within their intervention. Two studies were based on the Vermont Oxford Network (VON) approach(92), which is a slightly different from the Breakthrough Collaborative Series and therefore not considered in the analysis here. While three of the studies showed a positive effect, two of the studies (of which one was a RCT) did not show any significant effect. This review was similar to the findings by Schouten, but with the addition of one more study-that by Howard et al(122), looking at the effectiveness of Collaboratives in improving organ donation by eligible donors. It was a controlled before after study design and showed significant improvement in donor conversion rates.

In addition, Hulscher et all in their review also identified 67 articles that had an uncontrolled study design (93). The uncontrolled studies reported on improvements made in patient care and organization performance. These studies mainly looked at QI Collaboratives and process pathways while the controlled case studies (n=13) measured and reported both processes and outcome parameters. Studies with process of care parameters more often showed positive results than patient outcome parameters. The authors suggest that this might perhaps be because of the timeframe chosen. Shorter timeframe means that the more proximal process indicators are more likely to show effect while outcome indicators, especially for complex interventions such as QI, less so because it takes time for the improved processes to effect improvement in outcomes.

In summary, the evidence regarding the impact of Collaborative models on change in outcomes is very limited(20, 77, 80). But interestingly, some organisations have been able to achieve remarkable improvements using QIC methods while others have achieved modest results(86, 123). So the question is what is it in those organisations where the intervention worked that was different? According to Powell et al, where there is no improvement, it could be because organizations overall objective could sometimes be in conflict with the changes proposed by QI teams. Sometimes changes to one

unit could displace problems to another part of the system while at other times addressing certain aspects of care may expose other long standing problems within the organization that are difficult to address(77), Some organizations only make partial use of the PDSA cycles It has also been argued that PDSA are useful for small scale changes but unsuitable for system level changes(124). Engagement of the senior leadership of the organizations is sometimes difficult and this is especially true of clinicians(85, 125). Difference in perspective between professionals from different health care organizations can also be a challenge(124).

Hulscher et al try to review this apparent inconsistency in the evidence regarding the effectiveness of Collaboratives, by looking at the potential determinants of a successful QI Collaborative(100).

2.5.4 Determinants of successful QI Collaboratives

Several potential determinants of QI Collaborative have been suggested in literature(77, 100, 126, 127), that are necessary for a functional Collaborative. However, the evidence base supporting most of these determinants is either weak or non-existent. I have categorised the list based on the QI framework elaborated in Figure 9.

- Will-Ideas-Execution:

Leadership

Involvement of senior management and physicians(128) is important though the evidence on that is not well established(106, 129).

Alignment with organizational strategic goals.

Teams are more likely to be successful when the work they are engaged in is perceived as part of the organization's strategic goal(127, 130).

- Model for Improvement

Appropriate choice of topic

Choice of topic is considered an important determinant of QI Collaboratives. In general it was found that clinical projects were more successful than operational improvement projects(131) .

Ideas for improvement

Experts support for improvement is important(132). Facilitators are a huge resource to the Collaboratives and they provide intensive support to implement and maintain the QI interventions(81).

Preliminary work and learning

Improvement interventions are more likely to succeed when developed with, rather than imposed on, healthcare professionals.

Partial use of PDSAs

One of the challenges of QI Collaboratives has been to get the participating teams to fully implement the PDSA cycles. This does influence the success of Collaborative(86).

- Breakthrough Series/QI Collaborative

Data collection

Data infrastructure is an important aspect of the QI work and presence of good data systems does influence the success of QI interventions(106).

Characteristics of QI team

In general, support of QI from clinical practitioners is limited(133), but physician involvement does have a greater impact on the success of Collaboratives(134).

Other characteristics of the QI team include group climate i.e. group learning, orientation and group processes such as perceptions of feedback, task and relationship conflict(135).

- Context

Resources

The commitment of resources usually requires the support and commitment from the senior leadership. But at the same time there is also evidence to show that countries with highest spending on health are not necessarily the one with the best results(136). Along with availability of resources, the appropriate management of these resources has to be thought through and quality improvement techniques can be used to achieve this.

Organizational characteristics

The literature on the relationship between organizational culture and outcome defined in terms of quality of care, is very limited but does show a positive association with QI success(137-141). Other organizational characteristics such as organizational maturity i.e. the amount of time an organization has been formally involved in QI is also an important determinant of QI initiatives.

WHO has identified 6 domains necessary for the implementation of quality improvement interventions which includes leadership as the primary hub followed by information, patient and population engagement, regulation and standards, organizational capacity and models of care(142).

However, there is little empirical evidence to suggest these factors have an influence on the effectiveness of quality improvement interventions. The 23 papers reviewed by Hulscher et al(100), for instance, did not find any positive effects of leadership support, time and resources on QICs.

I have used the QI framework to enumerate the key characteristics of Collaboratives. An alternate taxonomy using Who-Where-How-What domains also highlighted similar Collaborative characteristics(94). There is also a general acceptance regarding the poor or practically non-existent empirical research on success factors of QI Collaboratives. For instance, while individual, organizational and clinical leadership have been identified as quite central for QI interventions (94, 100, 142), the actual evidence base supporting the role of leadership in QI Collaborative is very thin. The role of

leadership in fostering quality improvement is more of a shared perception among stakeholders than based on empirical data. This is especially reflected in studies that do not have comparison data (uncontrolled designs) and are largely based on participant's opinion of what's important.

A purely quantitative approach identifying each single factor, as done by Hulscher et al in determining the success of QI interventions, might not be the best approach to identify the key determinants. A successful QI Collaborative is likely to be a combination of few factors. The existing literature does not address how a collective action of these factors can be evaluated to determine success. Nor does it take into consideration the degree of influence these determinants have on each other. A quantitative approach ideally, would have weights attached to each success factor depending on various parameters, such as their degree of influence on other determinants. For instance organizational readiness has a bearing on QI team characteristics and on the preliminary work and learning that has to happen even before the implementation of a QI Collaborative. Current literature is unable to capture the inter-relationship between the various factors.

In summary, there is limited evidence available about the effectiveness of QI Collaboratives(48), Also, the available literature does not provide a clear picture as to what constitutes a successful QI Collaborative.. The limited or lack of evidence regarding effectiveness of Collaboratives arises partly from the ambiguity regarding what constitutes a QI interventions(84, 143)and partly because there is lack of clarity on what needs to be evaluated. A fundamental question related to this is about the choice of methods for evaluating QI Collaboratives(144). In areas where interventions have been reported to be successful, scaling up remains a challenge. There is no definitive or consistent methodology described in literature on how one can evaluate this scale up. There has been a demand in recent years for a robust evidence base for quality improvement Collaboratives, so as to assess their effectiveness among diverse settings, and develop insights into the specific processes and mechanisms by which they operate(66, 80).

2.5.5 Evaluating QI: Design and Methods

Using the Donabedian' structure-process-output as a heuristic framework, I begin with a discussion of evaluation tools and techniques for measuring inputs and processes, evaluating outcomes in terms of either quality of care itself or health care utilization or hard health outcomes such as mortality and morbidity. Different evaluation design and methods for measuring QI Collaboratives are discussed in this next section. Figure 13 presents an outline of these evaluation methods.

2.5.5.1 Evaluation of structure and process

A key feature of measuring the effect of an intervention is measuring its implementation strength. The steps often include identifying essential components to be measured, grouping components into domains, building a measurement instrument, piloting the instrument and finalizing the instrument(98). It can be measured by calculating scores for individual indicators and their domains, averaging domain scores or by weighting the components. Despite progress in recent years in measuring implementation strength, there is no consensus on defining or measuring implementation strength(98).

However, measuring implementation strength tells only part of the story. Unless one understand the interaction between the context & the mechanism by which the intervention is operating it will be difficult to attribute implementation strength to changes in outcome. This is discussed in detail in the discussion chapter (§8.2.6).

A tool that has been used in healthcare since 1960s and widely used in QI interventions since the mid 1990's(145) is the Statistical Process Control (SPC). Their strength lies in their ability to identify change that results in improvement-a central tenet of QI interventions. SPC has great value because it provides a quantitative measure of the implementation process and together with its primary tool –the annotated control chart provides a simple method for better understanding and communicating data to lay decision makers(73). It provides real time data on the process of care and through its feedback

mechanism, provides an evidence-based approach to systems management(70, 146). Within health care, it has been used in a wide range of settings and in over 20 specialties(70).

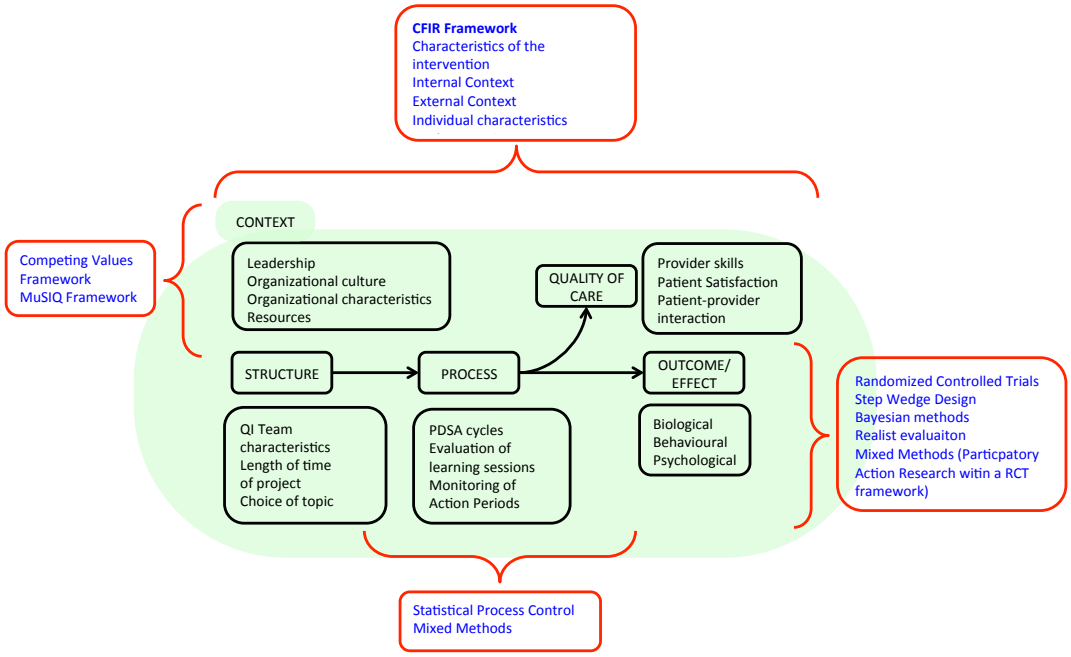

Figure 13: QI Evaluation methods

SPC can be used for improvement, monitoring as well as evaluation(70). For instance, Curran et al in their study on MRSA infection, demonstrate the use of SPC data both for improvement as well as for evaluation(147). Although in current practice it is not as widely considered for process evaluation, it does have the potential to provide valuable information on the 'black box' of the interventions and thus complement QI evaluation. However, it is not clear from literature if this method is applicable and appropriate for evaluation purposes. It be argued that SPC is similar to time series regression There is also a literature void in terms of detailing the possible implications of using improvement data for evaluation purposes(70, 148). While SPC can be very precise in measuring improvement in the process of care, linking this to outcomes can be difficult since outcomes are influenced by a host of other factors beyond the clinical microsystems level and as Brown and Lilford suggests it offers a low signal to noise ratio(149). However, SPC can be incorporated into integrated system measures (150) that attribute

improvements in quality of services to outcomes measures. Other techniques such as single case methodology have also been suggested(96).

Another useful tool is the CFIR [Consolidated Framework for Implementation Research](151). CFIR is a synthesis of various implementation theories that provides a pragmatic structure for assessment of the implementation context and for evaluating implementation progress. It is composed of 5 domains with 37 constructs(152). The five major domains are: intervention characteristics, outer setting, inner setting, individual characteristics and implementation process. Each of these domains consists of a set of constructs. There are 8 constructs related to the intervention characteristics, 4 constructs related to the outer setting, 12 constructs related to the inner setting, 5 constructs related to individual characteristics and 8 constructs related to process, adding up to a total of 37 constructs. These constructs are known to influence implementation. However, a major limitation of CFIR is that it does not specify the interaction between the constructs.

Since process evaluations focus on how an intervention is delivered rather than its effect on participants(153), process measures on their own are not sufficient. In order to know whether an intervention has had an effect on the outcome measure, research needs to integrate tools and techniques used to measure improvement process with conventional evaluation methods such as analytical or experimental study designs(154). Bayesian techniques which allows for integration of different types of evidence, allows for this type of synthesis(155).

2.5.5.2 Evaluation of outcomes

Various analytical designs have been suggested for the evaluation of outcomes(156). This includes RCT and step wedge designs, Bayesian methods and realist evaluation(155, 157, 158).

The pre-post test design is one of the most commonly used designs in the evaluation of QI interventions but has its limitations. It is prone to post-hoc reasoning and cannot rule out pre-existing trends. Most of the studies in

Schouten's(20) and Franco's(18) reviews were non-randomised and non-controlled studies. These study designs relied on post measurement, used before after study design without accounting for secular trends, used self-reported parameters rather than medical records, included anecdotal information and selected samples from selected sites. Such uncontrolled studies suffer from design limitations, are methodologically weak and probably biased in favour of positive findings.

Several quasi-experimental designs such as time-series, equivalent time-series, multiple-baseline study designs have all been suggested for evaluating QI interventions both as a single case as well as comparative group application(96). These designs are not without their limitations and their potential has not been demonstrated in mainstream QI research. For instance, one of the weaknesses for time-series design is its failure to attribute changes to the intervention(96, 159).

In the hierarchy of evidence(160), randomized controlled trials are considered the gold standard for evaluation(161). However, there are limited studies that have used a RCT design in the evaluation of QI interventions or more specifically in evaluation of QI Collaboratives and fewer still have obtained significant results. Homer et al evaluated the impact of a 12 month QI Collaborative (combined with Chronic Care Management on children with asthma (outcome measure)(112). They also analysed 3 process of care parameters. No overall effect of the intervention on the process or outcome measures were observed(112). The MERIT Study from Australia which used a cluster RCT design did not show an effect of using QI interventions on rapid response teams in ICUs (162). An RCT equivalence study by Gustafson and colleagues, on improving the effectiveness of addiction treatment services, showed that clinical-level coaching was equally effective as a combination of Collaborative components(159). A cluster randomized controlled trial evaluating the effectiveness of QI Collaboratives in improving colorectal cancer screening rates in primary care practice did not show any statistically significant improvements in screening rates between intervention and control

arms(163). The MaiKhanda trial looked at the effectiveness of a combined facility based QI Collaborative intervention and a community based women's group intervention and showed a 22% reduction in neonatal mortality(24). This is perhaps one of the first studies to use a RCT design for evaluation of a QI intervention within a low-income country.

All the studies covered a wide range of topics, different lengths of implementation period, have different levels for the unit of randomization and analysis, rendering any form of comparative analysis infeasible. Quality improvement interventions are generally classified as complex interventions. A major challenge in evaluation of complex QI interventions is that using rigorous scientific evaluation methods such as RCTs require interventions to be standardized whereas adaptability and change are inherent characteristics of such interventions.

However, Hawe et al (164) argue that complex interventions can apply rigorous evaluation techniques by standardizing the functions and process of the intervention rather than its components. In other words, the integrity (or fidelity) of the intervention is defined in terms of the evidence fitting with the theory or principles of the hypothesized intervention.

A major criticism of RCT is that they do not provide an understanding of how or why the intervention worked (158),(165). For instance, if the trial shows a positive effect it is difficult to judge on the generalizability of the trial findings without appropriate knowledge of the context and the process. For a trial with negative results, one is not sure if the intervention design was flawed or if the implementation was ineffective or if the results were as a result of inappropriate evaluation design(67).

Another limitation of RCTs evaluating complex interventions is that they are seldom generalizable given the rigour (where focus is on maximizing internal validity) applied in scientific studies. Random allocation reduces the degree of benefit that comes from an individual or group actively choosing to participate in the particular intervention. RCTs are not replicable especially in the QI

context where the direction of the intervention is greatly influenced by the character and composition of the QI teams and the organizational and other contextual factors(166). They are also relatively expensive.(167)

Given issues around generalizability, replicability and the relative expensive nature of RCTs, it has been debated whether it is an appropriate method for evaluating QI interventions(146, 167). Longitudinal time-series analysis using SPC have been proposed over RCT as an evaluation method for QI interventions(73, 167, 168).

Experimental evaluation, in general, tends to ignore the complexity of causality. Complex interventions are more than the sum of their parts and therefore one has to use a combination of evaluation design strategies to ascertain causality. Such a design also needs to consider multiple outcomes.

Supporters of realist evaluation methods, suggest a generative model of causality where context and mechanism interact to influence an outcome (as opposed to the successionist model, which underpins clinical trials, where "*cause X is switched on and effect Y follows"* (169).

Campbell et al propose a phased approach to developing randomized controlled trials of complex interventions using a combination of qualitative and quantitative evidence(30). This includes a pre-clinical or observational phase where theoretical concepts are clarified and modelling techniques used, an exploratory phase, followed by an explanatory phase which uses RCT design and finally a phase IV (or long term implementation phase).

Brown et al suggest a modified Donabedian causal chain model (structure-process-outcome) as a framework for evaluating complex interventions. They suggest a mixed methods approach (using quantitative and qualitative methods) to evaluate across different levels of the causal chain and then conduct a subjective quantitative synthesis of the diverse information using Bayesian methods(149, 170).

Mixed methods have been suggested by others as well(171),(29),(27) . A mixed method has several advantages. Leykum et al, suggest that using Participatory Action Research (PAR) within a RCT framework has the potential to take account of the local differences in healthcare systems and thereby allow some degree of generalizability.

Speroff et al advocate alternative experimental designs such as step wedge design(96, 170, 172). This is similar to a phased RCT, which could be the method of choice if it is not possible to implement an intervention simultaneously across an entire province or region, especially if there is limited or no evidence regarding the effectiveness of the proposed intervention.

Another method proposed is to develop an ex-post theory of what worked in cases where process evaluation was not fully conducted during programme implementation(67). This method, for instance, provided insight into the Michigan keystone project which very successfully reduced central line contamination in 103 participating ICUs in Michigan, using a Collaborative approach(173). However, using such theory-based approaches is prone to outcome bias since the theory is developed retrospectively after the outcome has been observed.

2.5.5.3 Evaluation of context:

There is conceptual ambiguity and methodological weakness in the current research of context in QI. There is also paucity of literature in understanding the inter-relationship of the contextual factors to each other and their effect on outcome(106).

The WHO perspective is that evaluating QI interventions cannot be independent of the context(142). Evaluating context involves evaluating a combination of summative and formative outcomes of intervention effectiveness. Organizational and system factors such as organizational and system capability, organizational culture, learning and social networks have been identified as drivers of change for large-scale improvement

interventions(174). The notion of (organizational) culture change as a key element of health systems improvement is gradually gaining wider acceptance(175). However most studies suggesting a link between organizational culture and health systems performance are methodologically weak(108) (106). A review of the quantitative instruments for the measurement of organizational culture in health care identified 13 instruments. The choice of instruments for investigation depends on the conceptualization of organizational culture by the research team. The assessment of organizational culture requires greater utilization of qualitative methods but, there is a predominance of quantitative methods used in literature(108). Various frameworks, such as the competing values framework(105), and the MuSIQ(109) framework have been suggested to evaluate organizational context.

Black and colleagues have suggested the development of a National Evaluation Platform (NEP) that will incorporate the broader context, by involving other sectors such as education and agriculture, and take measure factors such as women's literacy status, rainfall and agricultural output when measuring impact(176). NEP uses a dose response relationship where, dose is program implementation strength and response is increase in coverage and decrease in mortality.

The Kirkpatrick model can be adapted as a useful framework to link the changes in organizational context with outcomes(177). Outcomes are broadly categorised into reactions, learning, behaviour and results. It has been widely used for training interventions in the private sector, and has been adapted and adopted to classify and measure outcomes in healthcare improvement interventions (178).

2.6 QI Interventions: Challenges for evaluation

Mary Dixon-Woods et al identified ten challenges of QI interventions, categorized in 3 major categories: design and planning of improvement interventions; organizational context, culture and capacities; and, sustainability, spread and unintended consequences(179). Although these

challenges are for quality improvement in general, evaluation of QI also faces similar challenges.

❖ Design and planning of improvement interventions

One of the strengths of quality improvement interventions is their use of data for decision-making. This, however, requires will and support from senior leadership to invest in training of staff and establish a culture of accountability in the organization through the use of routine available data. At the same time QI interventions also require the skills of QI team members to collect, collate, analyse, present and use data routinely and consistently for decision-making. From an evaluation perspective, individual and collective behaviours within QI teams are likely to introduce variability in data quality over time and across sites . Behaviours are influenced by their data skills but also their belief and approach to evidence based decision making and therefore use of improvement data for evaluation purposes needs further scrutiny.

While interventions are usually directed at the organizational unit or microsystem, outcomes are usually measured at the individual product or patient level(156). The challenge for research here is to develop and evaluate a logic model that links the organizational and individual level outputs to the patient level outcomes.

• Organizational orientation and support

The concepts of quality improvement, especially within the health care system, are fairly new to developing country contexts and require significant investments in training and capacity building so as to embed them within health systems. Evaluation therefore needs to have a considerable lag phase in their design. Evaluation of QI interventions also requires to recognize organizations as political systems(180) and study of how innovative concepts such as Quality Improvement get embedded within such systems. Using theory based approaches, such as the theory of '*Diffusion of Innovation'* can help understand these mechanisms(181, 182).

- Health Systems Context

The origins of quality improvement in health and medical care are largely from the developed country setting including the United States, UK and Europe. The context of medical care provision in these settings is different from that in the developing country settings. There is little understanding of the influence of 'health systems readiness' in influencing quality of care in such settings. Context classification and evaluation methods are slowly emerging and being considered by the research community for inclusion within mainstream research designs.

- Leadership

The challenges to leadership are many-fold especially in developing country settings. It might be difficult to engage leaders because they are already dealing with a myriad of complex and competing clinical and organizational demands with limited resources, inadequate staffing and shortages of equipment(179). There is currently a gap in literature not only in quantifying leadership engagement but also in trying to understand the link between leadership and its effect on care outcomes.

- QI teams

Putting a multi-disciplinary QI team together poses a challenge given the hierarchy within the medical care system. Lack of sustainability of leadership within the QI team poses a significant threat to the functioning of the QI team. Understanding the critical role of the QI team as the agency for change is fundamental to the spread and sustainability of improvement efforts within the health system. Equally important is to understand the "*interaction between the emergent expressions of the agency and the dynamic elements of the context"(183)* Carl May-NPT).

❖ Sustainability, spread and unintended consequences

Sustainability and spread is usually threatened when there is an over reliance on individuals as is the case in most low and middle-income countries where

human resources are already stretched. Understanding the processes and mechanisms of spread therefore becomes more important than isolated study of individual behaviour change. It is also unusual for interventions to diffuse and spread on their own. Evaluations using theory-based approaches can use theories such as diffusion of innovation(181), normalization process theory(183) or Deming's theory of profound knowledge(46) o help understand mechanisms of intervention spread.

While improvement processes are focused on one aspect of care, other areas of care can get comprised. Having a sustained overall improvement effort is a challenge for implementers. The intervention should account for different needs and have a constant overview of the different moving parts of the intervention and their possible un-intended consequences and unexpected outcomes. Un-intended consequences and unexpected outcomes pose a challenge for prospective research methods as they are difficult to account for in the initial planning of the evaluation design. Evaluation methods thus need to be interactive and flexible to the changes in intervention design with evaluators and implementers working iteratively and in close collaboration with each other.

- Resources

Both human and material resources are a challenge. The challenge is both in terms of availability as well as in terms of appropriate and optimum management of these resources. Support from senior leadership is important for this. While a random allocation of sites might seem a desirable solution to balance out these resource challenges across sites, evaluation should be able to capture how resource availability (or non availability) can influence broader contextual factors (such as staff motivation), which can then indirectly influence programme outcomes.

- Time

QI interventions are demanding and require investment of time in terms of the programme theory, design, selecting appropriate measures and data

collection systems, and assessing organizational capacity. It is difficult to determine what amount of time is required for an intervention to reach a 'tipping point' when outcomes start becoming visible and measurable.

Interventions especially social interventions change over time(184). The greater the time lag between the intervention and intended outcome-the larger the latency period, the greater the chances of extraneous factors influencing the outcome and the more difficult it will be to attribute the intervention to the outcomes.[96]

- Ownership

One of the expected outcomes of using the Collaborative approach is for the local health facility teams to take 'ownership' of the project or event. This is also important for the sustainability of the interventions in the long run. However, this is one of the biggest challenges in improvement effort(175). There is also the constant tension between external evaluators and 'owners' of the intervention. While external evaluation is preferred, to maintain objectivity, evaluation designs are seldom able to capture the dynamic nature of complex interventions. Furthermore, while implementation data can make major contributions to evaluation, its interpretation can vary between evaluators and implementers.

2.7 Summary

Quality improvement as a concept has been adapted from the production and manufacturing industry and is now being extensively used in healthcare interventions.

The challenge for health systems today is to design knowledgeable healthcare systems(69) that maximise the alignment between the current best evidence and what can be achieved within the local context(185).

QI in health care is focused on improving the quality of healthcare service delivery. This is important and works well in settings where the basic infrastructure and systems are in place. However, in resource poor settings,

improvements in service delivery alone will not guarantee changes in outcomes. For improving quality of care in resource poor settings, a whole systems approach is required. A whole systems approach looks at improvements required within the different building blocks of the health system(142) including the interaction between these building blocks as well as engagement of the communities that the health system is serving.

There are different models for quality improvement in health care. The success of these models requires multiple approaches with strong leadership, a culture of participation, flexibility in implementation, adaptation to the local needs, and a feedback mechanism.

Various challenges exist to the design, implementation and evaluation of QI Collaboratives. There is a sense of urgency and excitement to improve care, which is understandable but the urge to favour action over evidence does have implications for quality of care(186). The evidence regarding the effectiveness of Collaboratives is thin. Analysis of the controlled studies in Hulscher's review had mixed results-while some studies reported positive results others did not and this was influenced by the choice of indicators. Various methods exist for evaluating quality improvement interventions, each with its own merits and challenges. The choice of evaluation design will depend on what within the structure-process-outcome (S-P-O) framework needs to be measured.

Measuring outcomes of complex QI interventions within a complex health system has limitations especially when it comes to attributing changes in outcome to the intervention. Evaluating quality improvement interventions is particularly challenging in resource poor settings. Improvements in service delivery, as measured by SPC (statistical process control) and other techniques, will not necessarily measure a change in patient outcomes, because of the low signal to noise ratio - all the other system level factors which influence the outcome, other than the improvement intervention, also need to be factored in the analysis.

Nevertheless, evaluation designs range from cluster RCTs to realist evaluation all of which have been debated in literature. A more pragmatic mixed methods approach has gained credence in the recent years.

Evaluation of the MaiKhanda project used a mixed methods approach where the outcomes of the intervention was evaluated using a cluster randomised controlled trial design while at the same time the intervention processes were measured using methods such as statistical process control charts as well as measuring determinants of quality of care.

The next chapter presents the trial result, i.e. the outcome of the intervention measured using the conventional cRCT design. I look at the effect of the QI Collaborative on the neonatal mortality rate at the population level and neonatal case fatality rate at the health facility level. In this Chapter, I also cover some of the limitations of the cRCT design.

Chapter 3 Trial Methods and Results

0 provides an overview of MaiKhanda's intervention and the evaluation design, which included the impact, process and economic evaluation (§1.8).

The impact evaluation was conducted using a cluster randomized trial design. As part of my thesis, I summarize here the impact evaluation of only the supply side intervention, namely the effect of QI Collaboratives on newborn mortality at the health facility and population level. This study design and the results from the QI arm of the trial are described in Section 3.1 and 3.2 of this chapter. Section 3.3 contains secondary analysis and interpretation of the trial results. In Section 3.4 I discuss the limitations of the impact evaluation study and provide an introduction to the process evaluation study complementing the impact evaluation, used in the MaiKhanda study.

3.1 Study Design

The study was conducted in the central districts of Malawi in Kasungu, Lilongwe and Salima covering a population of approximately 3 million. The choice of districts was based on the directive from the Ministry of Health (MoH) in consultation with the Reproductive Health Unit (RHU).

The hypothesis was that a quality improvement intervention defined by a breakthrough collaborative series and rapid action cycles (§1.5.2), would reduce the newborn mortality in the intervention population by 30%.

The unit of randomization was a health centre cluster and its catchment population. The trial design was a stratified two-by-two factorial design where the health facility clusters were randomized to receive quality improvement interventions or not and the health facility catchment population clusters were randomized to receive the women's group intervention or not. The trial design was stratified by the two interventions and then by district so that a balance of the number of intervention and control clusters was maintained in each of the districts.

The average population in each cluster was 4,113. Of the 95 facilities identified at the beginning of the study, 13 facilities were excluded and only 82 facilities were included in the study (Figure 14). Of the excluded, 9 were CEmOC facilities The major change was the exclusion of 18 dispensary clusters. These were health facilities that did not conduct any deliveries nor did they have any of the MaiKhanda interventions in their catchment areas. In summary, 61 health facilities and their catchment population clusters were included in the analysis. There were 15 health facility clusters with QI, 17 control clusters with neither QI nor women's groups, 15 population clusters with women's group intervention and 14 clusters that had QI in their health facilities and women's group intervention in their catchment population.

In order to measure the impact, a randomly selected representative population cohort of 310,000 was followed through a prospective surveillance system, where key informants selected from within the communities actively followed up pregnancies, births and maternal and newborn deaths over a two-year period. HSAs [vi] (Health Surveillance Assistants) from their respective catchment areas acted as supervisors to the key informants. Data collected by the key informants were reported to MaiKhanda evaluation team through the HSAs. All women of child-bearing age in this surveillance population were enrolled as participants in the study. Details of sample size calculations, the randomization process and key informant based surveillance systems can be found in the MaiKhanda report(33).

The health care providers were actively involved in the quality improvement intervention in the selected sites and therefore, it was not possible to mask the service providers to the intervention. Similarly, since patients and populations were subject to the care giving process, it was not possible to mask them to the intervention.

[vi] *Health Surveillance Assistants (HSAs) are part of the formal health system in Malawi. They act as the link between their catchment population and the health facility*

3.1.1 Primary outcomes

The primary outcomes included:

- Neonatal mortality rate (overall, early and late) and
- Perinatal mortality rate (stillbirths plus first week deaths). ICD-10(187) definitions of stillbirth, neonatal death and perinatal death were used in this study (Box 2).

Stillbirth: "A stillbirth or foetal death is a death prior to the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy; the death is indicated by the fact that after such separation the foetus does not breathe or show any other evidence of life, such as beating of the heart, pulsation of the umbilical cord or definite movement of voluntary muscles." The ICD-10 criteria for stillbirth was modified to include births after 28 completed weeks rather than 22 weeks.

Neonatal death: death within the first 28 days of an infant after "the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of the pregnancy, which after such separation, breathes or shows any other evidence of life, such as beating

Box 2:Definition for outcomes used in the study (ICD-10)

3.1.2 Secondary outcome measures

For the QI intervention, the secondary outcome measures included newborn case fatality rate, percentage of skilled attendants at birth, deliveries at facility, availability of obstetric care signal functions(188). Box 3 outlines the definition of secondary outcomes used in this study.

Newborn Case Fatality Rate **–** The number of newborn deaths divided by the number of live births at the facility.

Percentage of deliveries with Skilled Attendance – The percentage of deliveries at a Health Facility – it is assumed that all deliveries at a health facility are done by skilled clinical personnel such as a nurse or a clinical officer.

Practice of signal obstetric care functions – The availability and practice of signal functions gives an indication of the status of the facility. Health Centres are intended to be Basic Emergency Obstetric Care (BEmOC) facilities and qualify as such if they are able to provide all 6 BEmOC signal functions: Manual removal of placenta; Manual vacuum aspiration; vacuum extraction and breech deliveries; parenteral antibiotics; magnesium sulphate; oxytocic drugs (ergometrine or oxytocin). Hospitals are intended to be Comprehensive Emergency Obstetric Care (CEmOC) facilities and qualify as such if they are able to provide all 6 BEmOC signal functions *and* the 2 CEmOC signal functions: Caesarean Section and Blood Transfusion. The MaiKhanda facilities intervention did not specifically work on improving the availability of signal functions.

Box 3: Definition for secondary outcomes

3.1.3 Ethics

An evaluation proposal was submitted to the Ethic Board in Malawi-the National Health Sciences Research Committee (NHSRC) for approval. The study was approved in Malawi in 2007

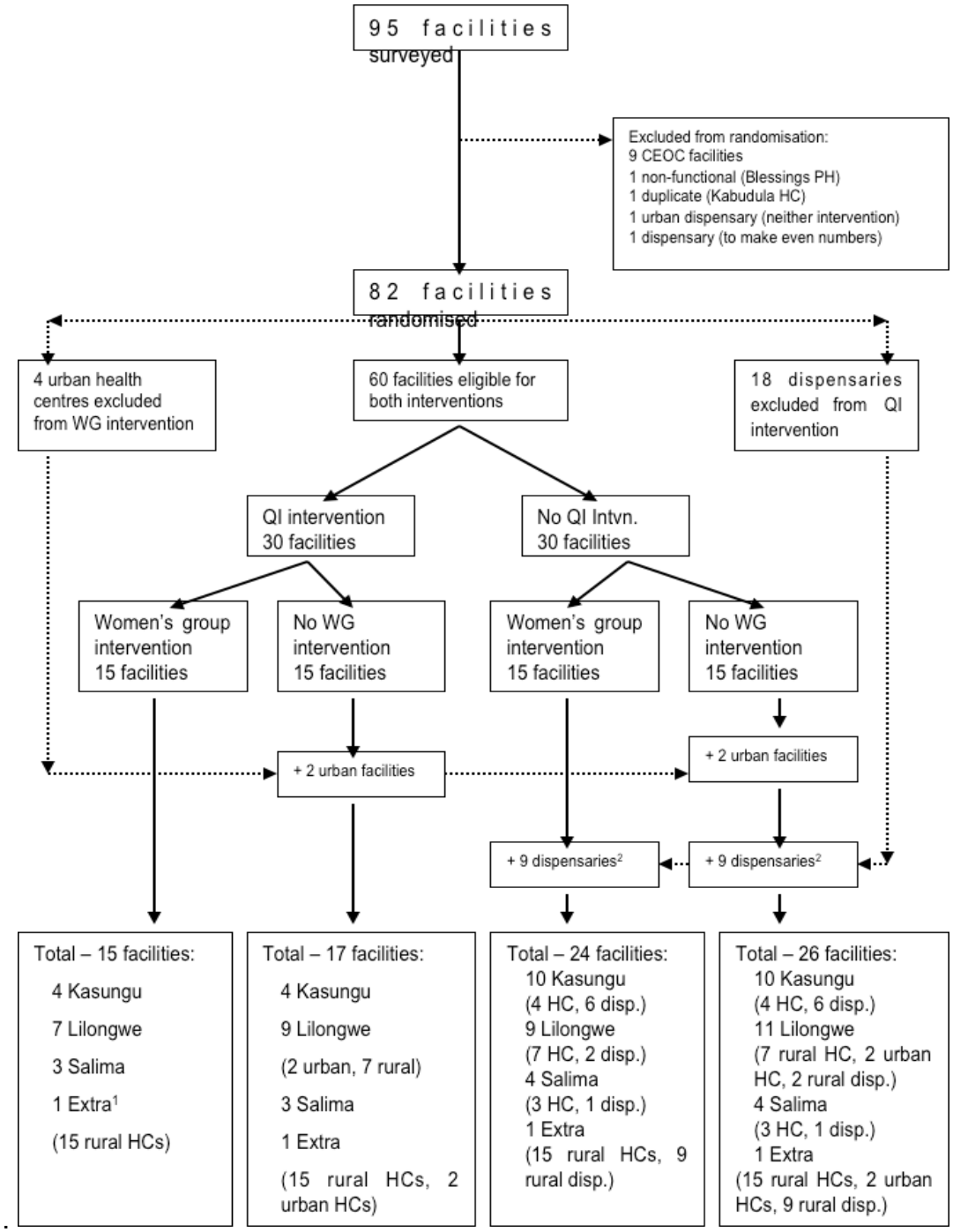

Figure 14: Trial Flowchart

3.1.4 Data quality

To ensure good quality data was collected, the following steps were taken:

- Training: HSAs were trained on the community surveillance form at the beginning of the project, with refresher training at meetings every month. KI were trained by HSA, with refresher training by MEO at quarterly meetings.
- Supervision: KI data collectors were supervised by HSA who were in turn supervised by MEO who were in turn supervised by UCL Technical Advisors (Bejoy Nambiar [BN] and Tim Colbourn [TC]) based in-country.
- Data entry: data was entered into a Microsoft Access database, with validation rules, by two trained data clerks.
- Data cleaning: duplicate records of women and babies were identified and removed from the database, and the assignment of all women and all deaths in the database to the correct cluster and district was checked and amended as necessary.
- Data verification: Deaths were verified by verbal autopsy interview with relatives of the mother or baby who died. A simple WHO approved algorithm based on movement, breathing and crying of the baby was used to differentiate between stillbirths and neonatal deaths.

3.1.5 Verbal autopsies

Verbal Autopsies were used to verify all deaths to ensure the quality of the data used in the analysis. Neonatal deaths and Stillbirths were followed up with a Perinatal Verbal Autopsy; The tools were translated into the local language (Chichewa).

Using a more advanced Bayesian algorithm, the InterVA method(189), it was possible to obtain causes of death for all 1817 PVA entered into the database. More details of verbal autopsies are available in the MaiKhanda Evaluation Report(33).

3.1.6 Analysis plan

The analysis plan was prepared in advance of the trial completion and analysis done in Stata 11.2 for Mac by IGH epidemiologist and Co-Investigator, Tim Colbourn. The analysis was a post-intervention comparison i.e. comparing mortality rates in the intervention and control group in the intervention period only. Differences in the intervention and control groups identified using the baseline data, was adjusted for potential confounding if they were found to be associated with mortality outcome. The analysis plan included a descriptive analysis followed by a cluster level analysis followed by individual level analysis.

Descriptive analysis showed time trends in baseline and pre-intervention period (Jun'07-Sep'08) and early and late intervention (Oct'08-Dec'10) period for the primary outcome measures namely neonatal and perinatal mortality.

Cluster level analysis was made of intervention effects on mortality by weighted t-test of cluster mortality rate summaries, taking stratification by district into account. Comparisons were made with and without other interventions, with and without adjustments for baseline mortality rates as well as looking at other potential confounders. The potential confounders include health facility deliveries, availability of signal functions within these facilities, nurse-delivery ratio at health centre, baseline staff psychology scores at health centre, geographic setting i.e. urban, peri-urban or rural, type of health facility (i.e. either government run or CHAM run facility), type of clusters (i.e. tobacco estates with high levels of migration or not). The effect of each of these variables on mortality rates was assessed first by descriptive univariable analysis followed by a step-wise logistic regression analysis of all variables.

The cluster level analyses were repeated at the individual level using logistic regression but as there was not data on the confounders at the individual level, adjustment were only made for cluster-level covariates.

An exploratory analysis of the trends in intervention effects was conducted by splitting the intervention period into two phases. The first phase was from Oct'08-Sep'09 and the second phase from Oct'09-Dec'10 during which the 'dosage' of the facility intervention was increased. There was general consensus from the implementation and evaluation teams that the original intervention inputs were inadequate to achieve the level of change desired in the health facilities and from October 2009 onwards the dosage of the facilities intervention was greatly increased.

3.2 Trial Results

The trial results are discussed in detail in the MaiKhanda's full evaluation report(33) and also in the subsequent publication(24). I was mainly involved in the interpretation of the analysis and subsequent drafting of the report and the paper. Here the summary result of only the QI intervention effect on newborn mortality is presented.

After cleaning, data quality was assessed using cluster summaries of the percentage of perinatal deaths without verbal autopsies (i.e. no verification of deaths). For the entire period of the study, 25.3% of the pregnancies did not register an outcome. However, in comparing the observed births with the expected crude birth rates in each of the clusters, the difference was about 3%, probably hinting at over-reporting of pregnancies by key informants. For baseline and intervention periods the loss to follow up was 19% and 29% respectively, with higher rates in the later months.

Around 14% of stillbirths and neonatal deaths remained un-verified. The percentage unverified did not differ significantly by RCT arm. Despite verbal autopsy, some stillbirths could not be categorized correctly as fresh or macerated stillbirth, or neonatal deaths as early or late. The percentage of uncategorized stillbirths or neonatal deaths did not differ significantly by RCT arm.

3.2.1 Baseline characteristics

Table 4 contains baseline characteristics of the facility and non-facility intervention areas. In general there is a good balance between the two arms of the trial except for skilled birth attendance, which is slightly higher in the intervention arm. In addition, the CHAM run facilities are slightly more in the non-intervention areas as compared to the intervention areas. These baseline characteristics were taken into consideration in the adjusted analysis.

Variables[a]		no FI	FI
number of clusters	n	29	32
Baseline NMR	mean	30.7	28.5
	se(mean)	2.5	2.9
Baseline PMR	mean	57.0	54.6
	se(mean)	4.6	3.6
Baseline Skilled Birth Attendance (%)	mean	45.5%	50.1%
	se(mean)	3.1%	3.6%
Deliveries per month per nurse at nearest HC	mean	21.2	25.2
	se(mean)	2.6	2.6
Signal Function availability at HC (number of functions)	mean	1.66	1.81
	se(mean)	0.24	0.28
Baseline staff psychology score (average of 19 questions on 1(=bad) - 5(=good) scale)	mean	2.40	2.46
	se(mean)	0.07	0.07
Urban (=1), Peri-urban (=0.5), Rural (=0) setting	mean	0.16	0.19
	se(mean)	0.06	0.06
Access by Tar Road (=1) or dirt road (=0)	mean	0.41	0.45
	se(mean)	0.09	0.09
CHAM (=1) or Government (=0) run	mean	0.22	0.14
	se(mean)	0.07	0.07
Tobacco Estates (=1) or not (=0)	mean	0.16	0.21
	se(mean)	0.07	0.08

Table 4: Baseline characteristics of non-facility and facility interventions (from MaiKhanda report)

3.2.2 Descriptive analysis

A total of 35,152 births were recorded from the pregnancy surveillance of which 1,059 (3%) were reported as stillbirths. Of the 34,093 live births, 16,828 were in the non-intervention areas and 17,265 were in the facility intervention areas. A total of 14,107 live births were recorded during the baseline period while 19,986 live births were reported in the intervention period.

Neonatal mortality rate during the baseline period was 29.6 per 1,000 live births. The NMR was slightly lower in the intervention arm (28.6) as compared to the non-intervention arm (30.5) at baseline (Table 5).

Outcomes	Baseline (Jun'07 - Sep'08)		Intervention (Oct'08 - Dec'10)	
	No FI	FI	No FI	FI
Births	7351	7225	9992	10584
Livebirths	**7114**	**6993**	**9714**	**10272**
Stillbirths	237	232	278	312
Macerated	44	42	58	49
Fresh	128	110	127	165
Neonatal Deaths	**217**	**200**	**310**	**284**
Early (0-6 days)	180	156	242	236
Late (7-28 days)	37	44	68	48
Perinatal Deaths	417	388	520	548
Stillbirth rate per 1000 births	32.2	32.1	27.8	29.5
Neonatal mortality rate per 1000 livebirths	**30.5**	**28.6**	**31.9**	**27.6**
Early neonatal mortality rate per 1000 livebirths (0-6 days)	25.3	22.3	24.9	23.0
Late neonatal mortality rate per 1000 livebirths (7-28 days)	5.2	6.3	7.0	4.7
Perinatal mortality rate per 1000 births (stillbirths and early neonatal deaths)	56.7	53.7	52.0	51.8

Table 5: Summary of births, deaths and mortality rates by FI and non-FI arm for baseline and intervention period (unadjusted for clustering, factorial nature of the trial).

Neonatal mortality during the intervention period was 29.8 per 1,000 live births. In the intervention arm it was 27.6 and in the non-intervention arm it was 31.9. Whilst there were fewer neonatal deaths in intervention areas (284) than non-intervention areas (310) during the intervention period (when the RCT analysis applies) there are actually more stillbirths (312 in FI areas compared to 278 in No FI areas), so when these are combined into perinatal mortality they cancel each other out.

The neonatal mortality remained the same within the baseline and intervention period, with the mortality going slightly up in the non-intervention arm in the intervention period. This is analysed further by looking at the variation in data by month and variation in individual data within clusters.

Figure 15 presents the neonatal mortality rate in the intervention and non-intervention arms by month. It is interesting to note that NMR appears to be going down in both the arms in the last 4 months of the intervention.

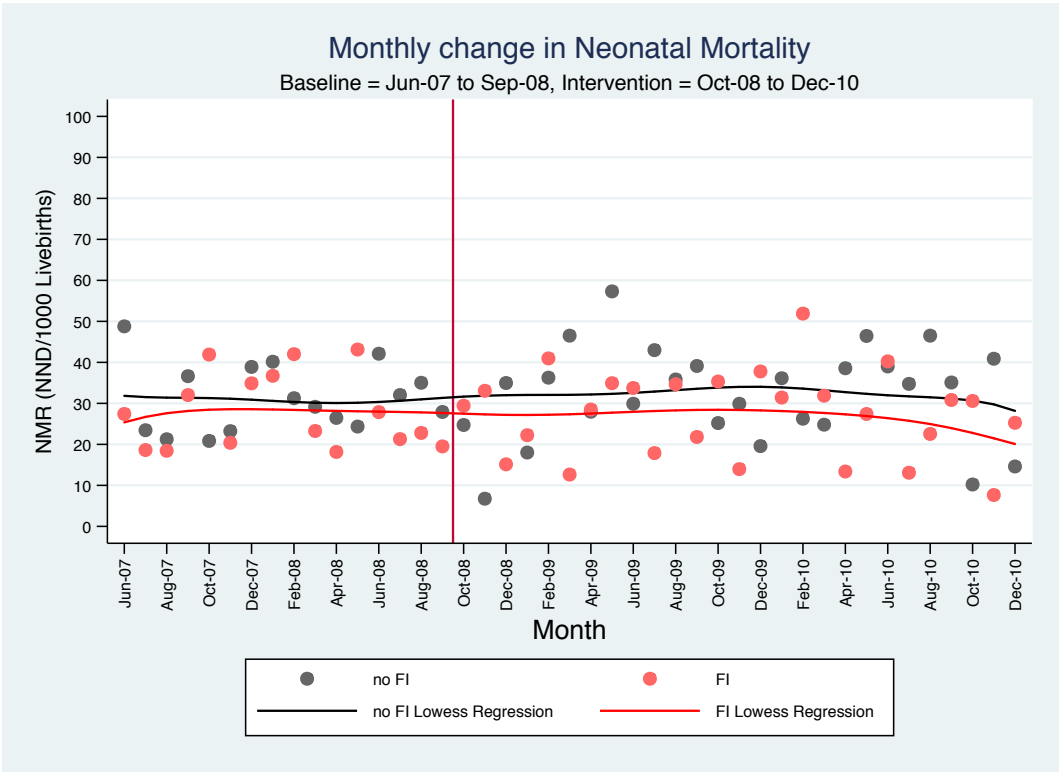

Figure 15: Neonatal mortality by month in the intervention and non-intervention arm (from MaiKhanda report)

3.2.3 Cluster analysis

The unadjusted cluster analysis shows a slight reduction of 4.3 neonatal deaths per 1,000 live births in the intervention clusters compared to non-intervention clusters, though not statistically significant (p=0.142). After adjusting for health facilities that were accessible by tarmac, the neonatal mortality showed a slightly smaller reduction of 3.6 neonatal deaths per 1,000 live births, also not statistically significant (p=0.213).

3.2.4 Individual level analysis

The factorial nature of trial was not taken into account in the cluster level analysis. This is accounted for in the individual level analysis, therefore making it more robust.

The individual level analysis of the intervention effect was done by calculating the odds ratio using logistic regression. This takes into account the effects of the districts (strata), stratification of the urban health centres and clustering by health facility catchment area as well as the two-by-two factorial nature of the trial.

The unadjusted analysis for FI interventions showed a 14% reduction in neonatal mortality [Odds Ratio=0.86 (95%C.I: 0.72, 1.03)], though not statistically significant (p=0.110). The adjusted analysis showed a 12% reduction in neonatal mortality [odds ratio = 0.88 (95%C.I: 0.74, 1.06)], also not statistically significant (p=0.176).

Intervention period (Oct 1 2008 - Dec 31 2010)		
Ouctomes	2 x 2 factorial trial	
	OR (95%CI)	p-value
Stillbirth rate per 1000 births	1.06 (0.84, 1.32)	0.638
Neonatal mortality rate per 1000 livebirths	**0.86 (0.72, 1.03)**	**0.103**
Early neonatal mortality rate per 1000 livebirths (0-6 days)	0.92 (0.75, 1.11)	0.373
Late neonatal mortality rate per 1000 livebirths (7-27 days)	0.67 (0.46, 0.97)	0.035
Perinatal mortality rate per 1000 births (all stillbirths)	0.99 (0.85, 1.15)	0.881

Table 6: Summary of RCT results

Table 6 presents the observed effects of the RCT on neonatal mortality[vii] and early and late neonatal mortality during the whole intervention period (Oct-08 to Dec-10). The effect of QI intervention on neonatal mortality is made up of effects on early neonatal mortality (Odds ratio 0.92; 95%CI: 0.75, 1.11) and late neonatal mortality (Odds ratio 0.67; 95%CI: 0.46, 0.97. The effects are significant for late neonatal mortality (upper 95%CI: 0.97, p=0.035).

Outcomes	1st Intervention period Oct 1 2008 - Sep 30 2009		2nd Intervention period Oct 1 2009 - Dec 31 2010	
	OR (95%CI)	*p-value*	*OR (95%CI)*	*p-value*
Stillbirth rate per 1000 births	1.18 (0.83, 1.69)	0.360	0.97 (0.77, 1.21)	0.767
Neonatal mortality rate per 1000 livebirths	**0.82 (0.63, 1.06)**	**0.124**	**0.90 (0.69, 1.17)**	**0.431**
Early neonatal mortality rate per 1000 livebirths (0-6 days)	0.92 (0.67, 1.26)	0.599	0.92 (0.70, 1.19)	0.516
Late neonatal mortality rate per 1000 livebirths (7-28 days)	0.54 (0.31, 0.92)	0.025	0.83 (0.49, 1.39)	0.473
Perinatal mortality rate per 1000 births (all stillbirths)	1.04 (0.83, 1.31)	0.719	0.94 (0.77, 1.15)	0.555

Table 7: Summary of RCT results by two intervention periods

[vii] *The intervention had no effect on perinatal mortality (OR=0.99) and hence not discussed here*

By further splitting the data into the two intervention periods, it is clear that the (non-significant) reductions in neonatal mortality in the first intervention period (OR=0.82) as well as during the whole intervention period (OR=0.86) is mediated largely by significant reduction in late neonatal mortality during the first intervention period (OR=0.54; 95% CI: 0.31, 0.92; p=0.025) (Table 7).

Figure 16 presents the picture graphically.

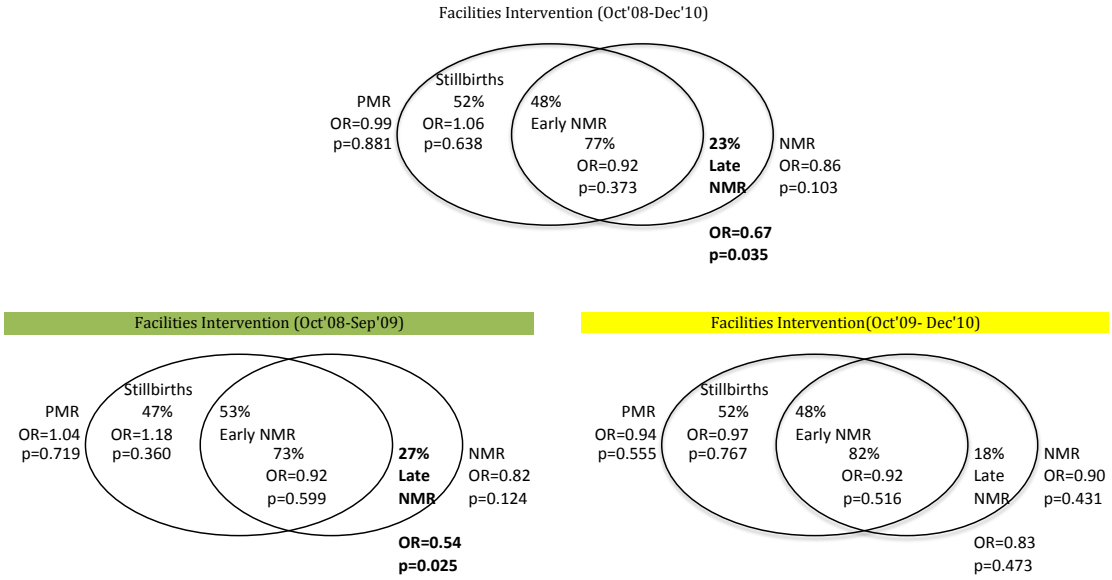

Figure 16: Relation between stillbirth, early and late neonatal mortality, by intervention periods.

Effects on late neonatal mortality in the first intervention period by the facility intervention are difficult to explain because interventions during this period were largely focused on the maternal change package (Table 2).

3.3 Secondary analysis

Since there was no statistically significant effect of the quality improvement interventions on neonatal mortality at the population level, a secondary analysis to see the effect of the intervention on neonatal case fatality rate at the health facility level was done. The analysis looked at the effect on NCFR (neonatal case fatality rate) both at the health centre level and at the CEmOC (Comprehensive Emergency Obstetric Care) level.

The health facility surveillance included data on the number and type of deliveries, maternal and neonatal complications, availability and use of EmOC signal functions and maternal and neonatal deaths and case-fatality rates. This data was collected on a monthly basis directly from the Health Centres by MaiKhanda staff, using a standard Health Facility Surveillance form ((33).

Once the forms were collected, the data was entered into a specially designed Microsoft Excel database that produced automated summaries and graphs of key data. Data were checked with reference to the paper forms and corrected where necessary.

Trend in facility case-fatality rates were assessed by several methods of time series analysis, including calculation of annual medians and means as well as quality improvement control chart statistics(33).

3.3.1.1 Health Centre

To begin with, there was an increase in the number of deliveries in all 3 districts for the period Jan 2007 to December 2010. The total number of deliveries in the health centres included in the study, went up from 28,359 in 2007 by about 13% to 32,121 in 2008/09 and by about 30% to 41,634 in 2009/10. The increase in number of deliveries in 2009/10 as compared to the baseline (2007) was about 47%. These findings are similar to the observations from the population surveillance data (Figure 22).

During the same periods the number of neonatal deaths were 133, 168 and 178 respectively. The neonatal case fatality rate was 4.8 per 1,000 live births in the baseline period and 5.2 and 4.4 per 1,000 in 2008/09 and 2009/10.

3.3.1.2 NCFR trends in the district

	Facility intervention			Control		
	2007/08	2008/09	2009/10	2007/08	2008/09	2009/10
KASUNGU						
Number of Health Centres	8	8	8	7	7	7
Deliveries	2,895	3,519	3,755	2,166	2,371	3,378
Neonatal Deaths	17	34	20	15	31	26
NCFR per 1000 deliveries	5.9	9.6	5.7	7.0	13.2	7.6
LILONGWE						
Number of Health Centres	15	15	15	19	19	19
Deliveries	9,673	9,617	13,387	8,098	9,745	12,884
Neonatal Deaths	45	41	59	24	39	45
NCFR per 1000 deliveries	4.9	4.3	4.4	3.0	4.0	3.5
SALIMA						
Number of Health Centres	6	6	6	7	7	7
Deliveries	2,493	3,139	3,760	3,034	3,730	4,470
Neonatal Deaths	16	13	12	16	10	16
NCFR per 1000 deliveries	6.4	4.1	3.5	5.3	2.7	3.7
TOTAL						
Number of Health Centres	29	29	29	33	33	33
Deliveries	15,061	16,275	20,902	13,298	15,846	20,732
Neonatal Deaths	78	88	91	55	80	87
NCFR per 1000 deliveries	5.7	6.0	4.5	5.1	6.6	4.9

Table 8: Neonatal case fatality rate in Health Centres, by district

In the 15 health centres in Kasungu, from a baseline of 6.4 per 1,000 deliveries in 2007, the NCFR rose substantially to 11.1 in 2008/09 before falling back to 6.6 in 2009/10, for all health facilities in Kasungu. When comparing intervention and control clusters for baseline and interventions periods in Kasungu, the NCFR is slightly lower in the intervention facilities (Table 8).

In the 34 health centres in Lilongwe, the NCFR remained fairly stable throughout the project at 4 per 1,000 deliveries. When comparing intervention and control clusters for baseline and interventions periods, the NCFR is slightly higher in the intervention facilities (Table 8).

In the 13 health centres in Salima district, the NCFR was 5.8 per 1,000 live births in the baseline period and decreased to 3.6 per 1,000 liver births by 2009/10. The NCFR was slightly higher in the intervention facilities in the baseline period but was about the same as in the control facilities in the intervention period.

Comparing across districts, the NCFR was similar in Lilongwe and Salima while Kasungu recorded a higher NCFR both during baseline and in the intervention period.

3.3.1.3 Comparing intervention and control areas

The number of deliveries was slightly lower in the non-intervention facilities as compared to the facility intervention sites, but there is a larger overall percentage increase in deliveries in the non-FI sites to eventually have an equal number of deliveries in both the sites towards the end of the intervention period.

There has been an increase in the absolute number of neonatal deaths both in the intervention and control areas, but looking at the neonatal case fatality rate, there is not much change in either of the sites. The overall neonatal case fatality rate in the intervention arm was 5.7, 6.0 and 4.5 per 1,000 deliveries for the three time periods respectively. The corresponding NCFR in the control arm was 5.1, 6.6 and 4.9 per 1,000 deliveries. The NCFR shows an initial increase in 2008/09 followed by a decline in 2009/10. These are not statistically significant changes.

In summary, while there has been an increase in deliveries in the health facilities over the period of the study, there were no real changes in the neonatal case fatality rates in the health facilities in the intervention and control areas over the period of the study.

3.3.1.4 Analysis of neonatal deaths in the CEmOCs

Although the CEmOCs were not part of the RCT evaluation, a considerable amount of QI activities were carried out in these facilities. However, most of the interventions in the CEmOCs were focused on maternal rather than newborn health, especially in the first intervention period (Table 2)

The CEmOC facilities were evaluated independent of the RCT. The analysis included trends in NCFR over the baseline and intervention time periods. It also included the use of run-charts, a standard method used in improvement measurement. The control charts show variation and assess the stability of

neonatal case fatality rates. A statistically significant change is accepted to have occurred if there are six consecutive values observed below the baseline median as per control chart rules. The median is then recalculated when the data indicates that there has been a change in performance according to control chart rules(190). Data was analysed as part of the improvement effort by the intervention team for the period from May 2006[viii] to December 2010.

Looking at the data for the 9 CEmOCs where the FI was focused (Figure 17 and Table 9), the NCFR during the May-December 2006 baseline[ix] averaged 29 (median 26), decreased slightly in 2007, averaging 27 (median 26), decreased substantially in 2008, averaging 21 (median 20), remained at this lower level in 2009, averaging 20 (median 20), and increased substantially in 2010, averaging 25 (median 23). This represents an overall decline of 14% using annual averages or of 10% using medians of the monthly data. However, it is difficult to attribute reduction in NCFR seen in 2008 and sustained in 2009 to MaiKhanda as there was little MaiKhanda neonatal work happening in 2008 (Figure 17) and although it started in 2009 was not up to full dosage until 2010. Secondly, in the absence of a control group it is difficult to ascertain if the observed reductions are attributable to the project interventions or a general secular trend for improved NCFR throughout Malawi. There was an attempt to collect data from non-participating facilities through an intern but getting data from the HMIS proved to be a challenge for the limited time period the intern was available and subsequently the plan was shelved.

Since the intervention has not been able to demonstrate significant changes in newborn deaths either at the population level or at the health facility level, the next level of investigation is to see if there has been a change in associated mortality figures such as the stillbirth rate, the early and late newborn mortality rate and the overall perinatal mortality rate.

[viii] *QI work in the CEmOC facilities started in 2006, earlier than in the trial sites.*

[ix] *Data for January to April 2006 is unreliable with data for Bwaila, which has one of the highest NCFR, also missing for these months, so it was not included in the baseline estimate.*

Since only neonatal deaths data were collected, it is not possible to analyse the NCFR as early or late or the perinatal case fatality rate. Both macerated and fresh stillbirths, however, were recorded in the health facility surveillance form. Fresh Still Birth Rate (FSBR) increased throughout 2006 and 2007 before remaining approximately the same in 2008 and then reducing in 2009 and 2010 to around the original baseline 2006 level(33). It is unclear whether the initial rise was due to better recording of data or whether, given the lack of process data on partograph use and management of obstructed labour, the decrease can be related to the MaiKhanda QI work.

3.3.1.5 Secular trends in neonatal health facility case fatality and population level mortality in Malawi

Although the health facility intervention did not observe a change in neonatal death rates in the population or in the facility, there is a probability that this was because of the background changes that were happening in the country around the period of intervention implementation. Analysing secular trends gives an indication if these changes could have possibly masked the effect of MaiKhanda FI interventions.

The DHS 2010 shows a slight increase in neonatal mortality from 27 per 1,000 live births [95%CI: 23,31] during 1999-2004 (DHS 2004) to 31 neonatal deaths per 1000 live births [95% CI: 28,35] during 2005-2010 (DHS 2010), although these increases have been statistically non-significant. The 2006 MICS (Multiple Indicator Cluster Survey) conducted for the period 2001-2006, observed a NMR of 33 per 1,000 live births [95% CI: 29,38]. Although statistically non-significant, it is suggestive of a general decline in NMR in Malawi during the period of MaiKhanda evaluation from 2007 to 2010. Recent estimates have suggested a nominal decline of around 3.3% in newborn mortality rates for the period 2000-2012 (12, 191)

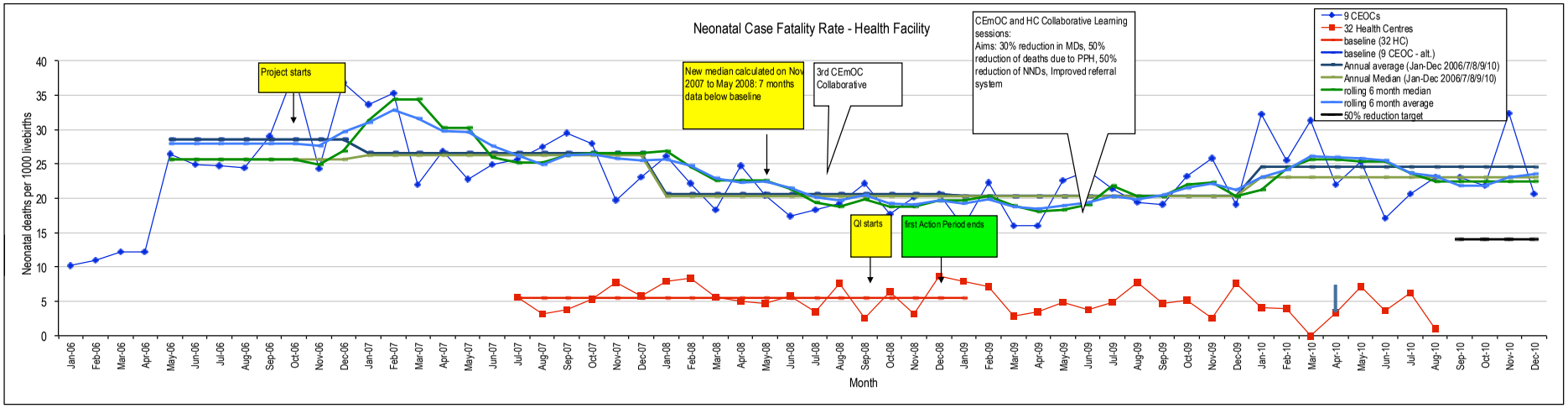

Figure 17: Neonatal case fatality rate at 9 CEmOCs combined, 2006-2010 (reproduced from MaiKhanda final evaluation report)

NCFR in MaiKhanda CEmOCs	May-Dec 2006	2007			2008	*annual % change*	Change since 2007	2009	*annual % change*	Change since 2007	2010	*annual % change*	Change since 2007
Livebirths 9 CEmOC	20,255	30,560			33,735	***10%***	**10%**	36,523	***8%***	**20%**	40,064	***10%***	**31%**
NND 9 CEmOCs	580	805			691	***-14%***	**-14%**	748	***8%***	**-7%**	981	***31%***	**22%**
NND KCH													
NND KCH assumed													
Annual average KCH assumed (rates higher?)													
	May-Dec 2006	2007	*change*	since 2006	2008	*change*	since 2006	2009	*change*	since 2006	2010	*change*	2006 baseline
Annual average (Jan-Dec 2006/7/8/9/10)	29	27	***-7%***	**-7%**	21	***-22%***	**-28%**	20	***-1%***	**-29%**	25	***21%***	**-14%**
Annual Median (Jan-Dec 2006/7/8/9/10)	26	26	***2%***	**2%**	20	***-23%***	**-21%**	20	***0%***	**-21%**	23	***13%***	**-10%**

Table 9: Annual Comparison of NCFRs in CEmOCs, 2006-2010

3.4 Making sense of the results

In summary, a cluster randomized trial looking at the effect of quality improvement interventions in health facilities in three central districts in Malawi did not detect a significant change in neonatal mortality rate at the population level. Further to this, a secondary analysis of the data from the health facilities was conducted which also showed no significant effect of the interventions on neonatal case fatality rates.

Absence of effect could be due to challenges in intervention design or in program implementation or challenges to the evaluation itself. A summary account of how the impact evaluation was implemented and the rigour applied in evaluation of the cRCT is presented in the section above (This has been summarized from the much detailed MaiKhanda evaluation report(33)). This current section covers the challenges in using RCT as an evaluation design for complex interventions such as QI Collaboratives.

RCTs are considered the gold standard in impact evaluation. The main advantage of this design is the availability of a randomized control cluster (or group) against which the interventions clusters can be compared, by controlling the context and other extraneous factors that can potentially influence the intervention attribution. Evaluation of the MaiKhanda intervention used a cluster RCT approach to analyse the impact of the intervention. But, the application of cRCT designs in the evaluation of QI (Quality Improvement) interventions poses many challenges both at the conceptual level as well as the operational level.

I briefly present the conceptual basis of MaiKhanda's improvement model so that the challenges to the use of cRCT design for impact evaluation can be better appreciated. At a conceptual level, the improvement model is based on Everett Roger's concept of "Diffusion of Innovation"(181) where the rapid spread of new ideas or practices happens largely by imitation(182) through individuals categorized as innovators, early adopters, early & late majority and laggards(192). The 'innovators' form a small proportion of this category but depending on their network linkages and inter-personal relationships, they can

engage the ‘early adopters’ who then go on to influence the ‘early majority’ and so on. The spread is thus organic and dictated largely by interpersonal network, which are not limited by the boundaries of intervention groups. Thus comparison groups can be considered as a hindrance to the organic spread of QI intervention.

In quality improvement interventions, the selections of groups or clusters (typically these are facilities within a health system) are influenced by their level of commitment or “readiness” to be part of the intervention. Such facilities are bound to possess individual and organizational characteristics that are inherently different from their comparison groups. Furthermore, comparison group in a cRCT are selected at the beginning of the intervention. The diffusion of innovation approach might be a more organic and perhaps pragmatic approach from an implementation perspective, but this conflicts with cRCT designs where the selection of a comparison group happens apriori.

From an evaluation perspective, focusing on only the ‘low hanging fruits’ does not tell us if the intervention has the same effect on the ‘not-so-enthusiastic’ facilities. A comparison group gives the opportunity to analyse and understand the mechanisms in play with ‘innovators and early adaptors’ as well as the ‘laggards’. However, from an improvement science perspective(46) psychology plays a role in the organic spread of QI interventions. Focusing on the low hanging fruit and managing to get a ‘critical mass’ of facilities that will adhere to QI principles, is a strategy that is more likely to influence the intervention acceptance by ‘later adapters & laggards’. This aspect of psychology of influencing a larger stakeholder group by building a ‘critical mass’ to eventually influence the outcome is difficult to measure using a comparison group, especially if the comparison groups have to be randomly allocated.

At an operational level, there are limitations to the identification and function of comparison clusters. The choice of comparison group is determined by the level at which the intervention takes place as well as the level at which

outcomes are measured. QI interventions can focus at different levels such as improving the processes of care or service delivery mechanism or re-organization of the healthcare systems. In most cases it is a combination of these different levels. These different levels are inter-dependent, making it difficult to isolate intervention effects even in the presence of a comparison group. Randomization at a higher level (e.g. district or organizational level) offers a better comparison cluster but randomization at this level will allow for very few clusters (there are only 28 districts in Malawi and it was not possible within the scope of interventions such as MaiKhanda to be active in all the districts), invalidating the results of the analysis(193). While having a comparator cluster at a much lower level such as individual service provider level, raises challenges for an appropriate comparative sample for outcome measurement.

Another important operational challenge in using RCT design is the unit of randomization itself i.e. determining the appropriate level of randomization for a QI intervention. In the MaiKhanda study randomization was at the health centre level whereas tertiary care and secondary care hospitals across the intervention sites also formed part of the QI intervention design. In fact, a significant proportion of the programme implementation efforts, especially in the earlier part of the intervention, were focused on the secondary and tertiary care (CEmOC) health facilities, which were not included in the trial clusters (Table 2).

As such Health Centres contributed to a smaller proportion of the mortality (around 16%) since most of the complicated cases (and therefore deaths) were referred to a higher level of care(33). Moreover, the study design considered all HCs as BEmOC sites, based on understanding from the baseline report whereas this was not the case(33). This was the intent of the Ministry of Health (MoH) but there were facilities that were not designated as BEmOC sites. But even where designated, they were out of stock or had insufficient human resources and therefore with limited functional

capacity. Thus having health centres as the unit of randomization itself could influence the results of the intervention.

The inclusion of CEmOC facilities in the intervention design could have influenced the trial results. Analysis of the CEmOC neonatal mortality data indicates that only 19% of all the neonatal deaths observed throughout the intervention period of the RCT occurred in the non-randomized CEmOC facilities(33). CEmOC work influenced the intervention and control arm equally. From an intervention design perspective, CEmOCs were a part of the intervention because the implementation focus was across a vertical slice of the health system.

The QI interventions were being introduced within the health system, which is complex and adaptive and composed of inter-dependent building blocks and levels. Changes introduced in one level of the system has implications and effects on other levels, manifested through feedback loop mechanisms. It is this inter-dependency of the different levels of the system which enable and constrain change and thus improvement(194). However, the extent of this dependency and its subsequent effect on newborn outcomes cannot be explained thru the RCT design.

Furthermore, given the inter-dependent levels of the health system, randomization at the health facility level will mean that comparison facilities will be influenced greatly by the intervention, especially since other levels within the health systems such as CEmOCs were also part of the intervention. It is not possible in complex social interventions, such as QI Collaboratives, to double-blind the participants and therefore difficult for trial designs have to take account of the biases that can be expected in such interventions. The design is also prone to ‘contamination’ effects as a result of these(28) Measuring the extent of this contamination between intervention and comparison groups is a difficult exercise and does influence the study results.

The next challenge is that of intervention and implementation fidelity. For complex QI interventions that have long implementation periods, comparison

sites are prone to be influenced by the intervention (referred to as 'spill-over effect') and contamination by other interventions cannot be ruled out(35).

A key principle of impact evaluation is to reduce bias by standardizing the interventions. In conventional experimental evaluation standardization of interventions is the norm. This means feedback loops are absent and there is no mid-course correction. However, these are important characteristics of QI interventions especially when embedding these interventions within the health system. It has been argued that it is still possible to account for the changes, if RCT designs can focus on maintaining the fidelity of the function (or principles) of the intervention rather than standardizing the activities (referred to as form of the intervention)(164).

Improvement models assume that an intervention works and if it does not work, incremental adjustments are made through feedback loop mechanisms, thus promoting continuous improvement. Thus PDSA cycles can be iterative leading to emergent properties of an intervention. It is not very clear how these emergent properties can be accounted for within a cRCT evaluation design. For instance, data improvement was an important activity of the QI intervention and while some facilities absorbed it completely, others were less enthusiastic about it. In areas where this was not completely implemented, it can provide a false picture of the changes occurring at the facility. For instance, with the increase in neonatal case fatality rates in health facilities in Kasungu district, it is difficult to conclude if this is a genuine increase or if this is a result of better recording of data. Since the data improvement is a continuous exercise, it is difficult to determine at what time point improvement kicked in and subsequently at what point the data should be considered as of credible quality.

3.5 Summary

The trial brought some interesting findings to the fore. Despite prescribing to a well-established improvement model, the intervention failed to measure an impact on newborn mortality at facility level and at population level. The challenges to measuring impact of complex interventions such as QI

Collaboratives is explored in this Chapter and it explains the limitations to the choice and application of evaluation design than actual implementation of the cRCT itself.

The trial results were probably influenced by the choice of intervention clusters (random v/s organic diffusion approach), choice of control group, (lack of) measurement of the influence of context, and intervention fidelity i.e standardization versus a feedback and iteration process that leads to a dynamic intervention model.

Process evaluation in the MaiKhanda evaluation report highlighted the contextual factors such as severe lack of material and human resources, lack of clinical skills, lack of motivation and morale, lack of interest or ability to use quality improvement methods or a combination of these factors. The PhD thesis goes further to explore how all these various factors, individually and combined, influenced the intervention outcome.

Evaluation also concluded that lack of observed effects could perhaps be due to sub-optimal implementation. This is the subject of further exploration in my PhD.

Given the complexity of QI interventions and the challenges of evaluating context and intervention fidelity, there is a growing consensus among improvement researchers to consider plausibility designs alongside probability designs(27, 29, 171, 195). In the next Chapter I explore the reasons as to why we did not see an effect, through the application of a plausibility design. I use a theory based approach to evaluation to try and answer 2 questions:

- Did the MaiKhanda intervention do the right things? This explores the characteristics of the intervention and the factors that might have influenced the results
- Did the MaiKhanda intervention do things right? This will explore issues related to the dosage and implementation of the intervention.

Chapter 4 Research Strategy

4.1 Introduction

Impact evaluation of the MaiKhanda intervention followed a linear logic model of evaluation design, analysing the input, process and outputs and finally the impact using a cluster randomised trial design. Results of a cluster RCT measuring the effect of QI interventions on neonatal mortality at population level did not show an effect (§3.2.4).

As part of the process evaluation of the MaiKhanda programme, the specific activities of QI intervention, and the key determinants influencing quality of care were explored. An attempt was made to triangulate the RCT results with the results of the process evaluation studies to draw inferences regarding the effect of MaiKhanda interventions. However, this did not go as far as explaining the context within which the interventions were implemented or the mechanisms and their possible influences on the intervention.

A theory based approach to evaluation can dwell into the black box of interventions and shed some light on the intervention mechanism. This theory-based approach to evaluation is the one that I have considered for my PhD. The approach involves drawing up a comprehensive research strategy to map out the causal chain from inputs to outcomes and impact and testing the underlying assumptions(35) and finally analysing the context and mechanism(196) by which quality improvement was delivered (or not) in the intervention.

Before explaining the theory-based approaches, a brief background of the programme pathway that is common across most QI interventions is presented here. An intervention begins with a set of activities including intervention design, implementation strategies, assumptions and risks which form an integral part of its implementation framework(197). Once an intervention is implemented with the appropriate implementation strength, (characterized by the dose, duration, intensity specificity and fidelity of the

intervention), it stimulates a change in thinking or behaviour of the agency (i.e individual and collective reasoning of all key stakeholders involved in the project) which in the appropriate context contributes to the success (or not) of the programme(158). This is referred to as the 'mechanism' of an intervention. This is the key point differentiating a complex intervention such as QI from clinical trials.

Thus in evaluating interventions using theory-based approach, there are 2 major aspects of the programme to consider. First is to hypothesize and analyse links between programme activities and its anticipated outcomes. This is referred to as implementation theory, prescriptive theory, 'little t' theory and in programme management literature as programme theory(198). I use the term *implementation theory* here to avoid confusion.

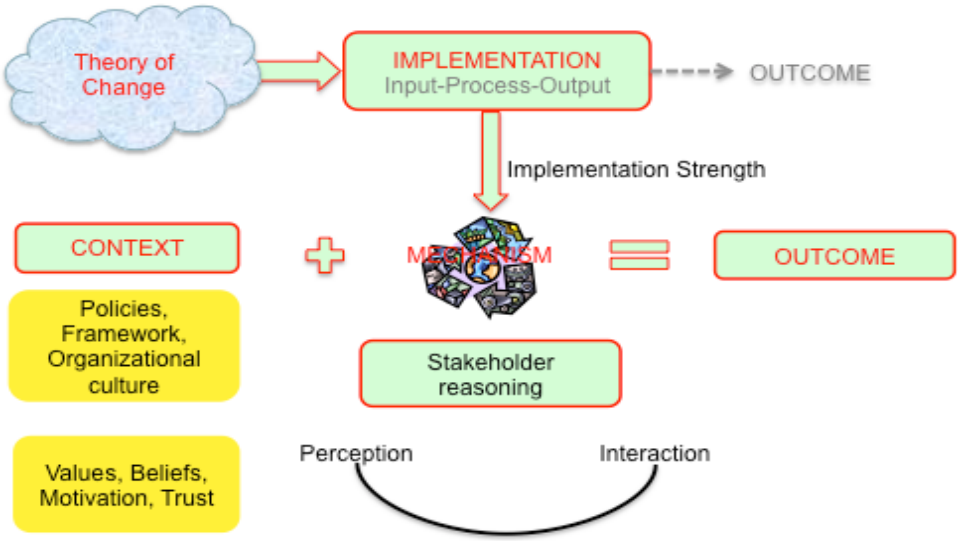

Figure 18: Overview of Theory Based Evaluation (TBE) approaches

Secondly, there is the causal links between the mechanism released by an intervention and its anticipated outcomes. This is referred to as programme theory, descriptive theory, big T theory and also as middle range theory in realist evaluation. I use the term *programme theory* here.

There are two main approaches to theory-based evaluation (among others) namely, the Theory of Change (ToC) approach and Realist Evaluation (RE) approach, and they take into consideration both the implementation theory and programme theory (Figure 18). Theory of Change has a strong base in explaining implementation theory but is less directive when it comes to explaining programme theory. Theory of change is analysed through development and testing of logic models(199) and is usually linear in nature. Here programme theory is analysed mainly in relation to its implementation strength.

Realist evaluation on the other hand is concerned largely with uncovering programme theory, usually using a C-M-O [Context-Mechanism-Outcome] configuration [x] (200). Here programme theory refers to the programme mechanisms (functions and strategies), outcomes the programme intended to generate and context that contributed to particular outcome(201). Realist evaluation tends to be contend with the theoretical explanation (rather than any empirical observation) of implementation process(67). Therefore there is less emphasis on the strength of implementation process within a realist evaluation.

But in reality, for interventions that are being tested in new settings, implementation strength is an important precursor to the CMO configuration framework. Thus a C-I-M-O is perhaps a more appropriate configuration, where I stands for 'Implementation'. I discuss this in detail in Chapter 8 (§8.2.5).

The challenge with complex social intervention is that it is difficult to predict the implementation strength ie the dose, duration, specificity and intensity of the intervention that is required to "trigger" the mechanism of the intervention(202). Nevertheless, a good understanding of implementation theory is important in order to conduct a realist evaluation(203).

---

[x] *A C-M-O configuration is a systematic description of the context, mechanism and outcomes and also the generative causal relationship between these elements*

The evaluation strategy for my PhD involves integrating the principles underlying Theory of change and realist evaluation approaches to evaluate implementation theory and programme theory respectively. My argument for combining both approaches is that a comprehensive model is required to provide a plausible explanation of the mechanism of the intervention(204).

The theory of change includes a broader theoretical framework followed by a conceptual framework. The broader theoretical framework explains the theoretical underpinnings of the IHI model for quality improvement used in MaiKhanda study (Figure 9. This entails Deming's system of profound knowledge and has been covered in some detail in Chapter 2 (§2.3). The conceptual framework is an operational version of the broad theoretical framework and includes the intervention design, implementation strategies, context, stakeholder engagement and both implicit and explicit assumptions and risks. The implementation strategies can be further classified into 6 major areas: organization of work, capacity building, finance strategy, restructuring, quality management strategies, attention to policy context (making programme sustainable)(205). The context refers to both the internal and external context.

The next phase of evaluation involves evaluating the programme implementation mechanism (also referred to as the programme theory). Here, a wider health systems research (HSR) approach is adopted where implementation, rather than being focused just on service delivery, is considered as an organizational, social & political process involving stakeholder perceptions (both individual and collective) and their interactions(206). These stakeholder perceptions and their interactions are guided by the hardware such as legal and policy framework within the health system while the software includes beliefs, value systems of individuals (Figure 18).

Stakeholder perception and their interactions form the basis for stakeholders' (individual and collective) reasoning which along with the context influences the mechanism of an intervention. The context includes structural factors such

as availability of drugs, human and other resources and functional factors such as organizational readiness and organizational culture. It also involves broader health systems environment and policy context.

The above described context and mechanism are an integral part of any intervention and does have an influence on its outcome. I attempt to explain this based on the principles of realist evaluation approach(200). A hallmark of theory based evaluation is its multi-disciplinary nature (and thus different from methods based research) and use of mixed methods. It uses a research strategy rather than a particular research method(206).

4.2 Overview of the research strategy

Using theory based approaches to evaluation, I am trying to answer two main questions:

- Did the intervention do the right things?
- Did the intervention do things right?

The programme evaluation uses a mixed methods approach for analysis. This includes qualitative and quantitative methods, contextual analysis and narrative documents.

The keys steps in my research strategy are outlined below:

1. It begins by conceptualizing the theory of change of the programme interventions (including the risks and assumptions) and its intended outcomes. This is detailed in § Chapter 5
2. Next step involves a collating and summarizing the available evidence from the MaiKhanda programme. This includes individual process evaluation studies such as staff motivation survey, resource availability survey, provider knowledge survey, women friendly care survey and the CEmOC survey (§Table 12) as well as summaries of relevant project documents such as the original proposal, the strategic review document and the 'data deep dive' report along with quarterly programme reports and Programme Management Board (PMB) meeting minutes. Field visit

observations from trip reports of senior technical staff are also included in the documentation analysis. I was involved in the design, management and analysis of all of the surveys as part of a research team and these various studies act as a secondary data source for my PhD analysis.

3. The next step involves synthesizing the multiple and varied strands of evidence into a framework for analysis. I use the CFIR (Consolidated Framework for Implementation Research) framework(151) to outline the domains and constructs, to consider for my evaluation and re-organize the available evidence according to these domains and constructs. This involves consolidating all the collated evidence and then synthesizing them for each of the individual domains and their constructs (§Table 12). This gives a broad understanding of the key components of the intervention and also helps to shed some light on the implementation strength of the programme. Chapter 6 provides a synthesis of all the available data from MaiKhanda project.
4. The interaction between the key components of an intervention and the underlying mechanisms it triggers off is analysed through an adaptation of 'realist evaluation' approach. Here, instead of developing new C-M-O configurations, existing mechanisms of other successful QI Collaboratives are compared with that of MaiKhanda. This provides an opportunity to refine the programme theory for quality improvement interventions. This is covered in Chapter 7 (§7.1)
5. The final step is to reflect back on the implementation theory, to try and analyse how or (why) the interventions did (not) trigger a mechanism. However, this cannot be done in isolation. It includes taking into consideration the implementation strength and the context within which it is being implemented. This is covered in Chapter 7 (§0)

4.3 Research Strategy

4.3.1 STEP 1: MaiKhanda's Theory of Change for its quality improvement intervention.

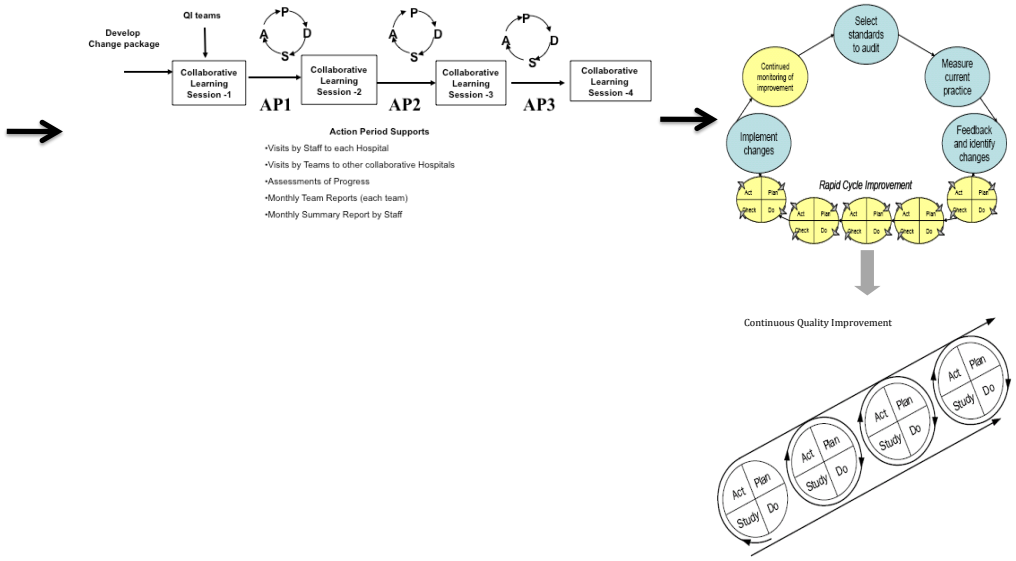

Figure 19: MaiKhanda Theory of Change (QI)

The MaiKhanda interventions were based on the three delays model by Thaddeus and Maine(13) described in the introductory Chapter (Figure 2). The quality improvement aspect of the intervention was based on the model for improvement developed and implemented by IHI(82), which can trace it origins to the improvement principles discussed in Deming's system of profound knowledge.

The method used to describe MaiKhanda's Theory of Change (Figure 19) is a narrative description from the original proposal document, the mid-term strategic review document, the baseline survey conducted in 2006 and the baseline retrospective evaluation conducted in 2007. The original proposal and the mid-term strategic review document gives an understanding of the intervention designs as well as helps define the implementation strategy. The 2006 baseline survey provides an overview of the context in Malawi in relation to MNH (Maternal and Newborn Health) at the beginning of the programme.

The 2007 baseline evaluation study provides details on intervention implementation at the beginning of the study i.e in the pre-intervention period.

4.3.2 STEP 2: Collating and summarizing the evidence

The various process evaluation studies related to quality improvement intervention that were conducted as part of MaiKhanda programme, are compiled and presented in Table 12. Using a quality of care framework(31), MaiKhanda's process evaluation looked at patient and provider components and resource availability. This includes the health facility resources survey, the staff motivation survey, CEmOC facility survey and provider knowledge study (§1.8.2). A detailed description of each of these studies is available in the MaiKhanda report(33). In addition, the process evaluation also included a case study on newborn resuscitation at St. Gabriel Hospital that was not included in MaiKhanda's final evaluation report ((§Appendix-2.7). Various programme related documents were generated by the implementation team such as the progress report, strategy document, PMB (Programme Management Board) meeting minutes, reflective evaluation (known as Data Deep Dive exercise) and meeting notes. These also constituted as evidence for implementation and program theories and were summarized directly into the CFIR framework. All this collated data is organized as a spread sheet and presented in Table 12

4.3.3 STEP 3: Consolidating and synthesizing the evidence

The Consolidated Framework for Implementation Research (CFIR) developed by Damschroder and colleagues(151) is a framework for formative assessment of intervention effectiveness within a specific context. It provides an organized set of 5 major domains (the intervention characteristics, the outer setting, the inner setting, the characteristics of the individuals and the implementation process) and 37 constructs, which can help, hypothesize theories regarding the mechanism of change, which can then be tested empirically. While this is a long list of constructs, the intention is not to evaluate all of them rather, analyse key constructs in the context of the study and that, which will be most valuable to the study.

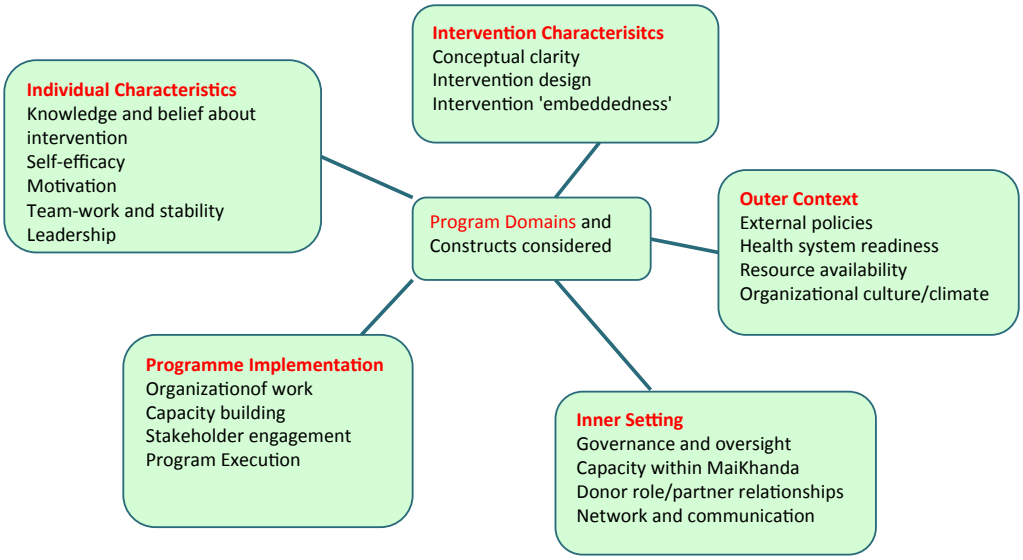

Figure 20: Programme Domains and Constructs considered for evaluation

From a realist evaluation perspective, the choice of constructs would be guided by the middle range theory that would be developed as part of the theory development process. But since I was evaluating retrospectively and from existing data sources, the selection of the various constructs was guided largely by the availability of data at hand rather than the ones most valuable to the study. This poses some limitation to the study. Firstly, it is likely that the constructs developed on the basis of available data might not overlap with constructs developed from drawing up a MRT. Secondly there has been no weightage given to the various constructs that might or might not overlap, it is difficult to determine their influence on intervention effectiveness. For instance, one important aspect of the intervention that resonates throughout the study is the role of leadership, both within MaiKhanda and within the health care delivery system. This was also identified by Hulscher and Schouten as a key determinant of QI success(100). While every attempt is made to explain this in the analysis section, it is limited by the paucity of data on this variable. Thus, while all the domains within the CFIR framework have been covered to describe the QI intervention of MaiKhanda programme, the selection of constructs is limited by the availability of data. This is one of the limitations of the study. The selected constructs are presented in Figure 20

All the evidence generated through the secondary data from various process evaluation studies and programme documents provide information for the constructs listed in Figure 20. By abstracting the evidence from the individual studies as well as programme related documents, a general description for each of the constructs is developed.

The programme has been divided into 4 phases, for the purposes of evaluation: the baseline period, the pre-intervention period, early intervention period and the late intervention period (§1.6). For each of the constructs, I attempt to synthesize the change processes during each phase of the intervention, from the consolidated evidence summaries in order to sketch a picture of the programme evolution.

4.3.4 STEP 4: Evaluating programme theory of MaiKhanda QI intervention

Next, I try to construct the mechanism of the intervention, using realist evaluation approaches. The crux of realist evaluation lies in understanding how mechanism(s) are triggered in a particular context and not others.

There are five major steps in the design and implementation of a realist evaluation. The first step is to elicit the programme theory and from that abstract a middle range theory (MRT). The other steps include research design and data collection tools and finally data analysis and synthesis. The research design and choice of tools is determined largely, by the key thematic areas identified from the middle range theory. The data analysis in realist evaluation adapts a 'thematic analysis' approach similar to ones used in social sciences, to abstract data from the various data collected.

Ideally, using a realist approach, especially for a retrospective study, one would develop a MRT (Middle Range Theory), as a first step. A MRT is defined as a level of abstraction that one is able to derive from the programme theory, but close enough to observed data such that it can be empirically tested"*(200)* (201). According to TBE, for an intervention there are several mechanisms, which are functional and may interact with each other in complex ways. These need to be considered simultaneously sometimes, to

make an understanding of the social phenomenon(207). Thus MRT is usually applied to separate the forest from the trees. Since analysing the context can be overwhelming, it has to be narrowed down to a level, where it can contribute meaningfully to refining the theory. Similarly, when dealing with complex adaptive systems such as the health systems, there is a need to focus on certain aspects of the intervention mechanism. Middle range theory helps achieve this objective. However this can also be a limitation for analysing programme theory. Thus a major drawback of MRT is that it is prone to 'tunnel vision' of the intervention as the theories are defined by the researcher and there are no set guidelines on how narrow the MRT needs to be. MRT is also drawn up in cases where the realist evaluation is retrospective and there is limited data (in terms of programme strategies and functions) available to draw up a CMO configuration

Within a realist evaluation framework, I considered a modified approach to elicit MaiKhanda's program theory, for the following reasons:

MRT is often used in realist evaluation to guide the refinement of proposed theories and identify the relevant constructs to evaluate them. Firstly given the process evaluation data and all the documentary evidence available from the MaiKhanda evaluation, I could extract data from the constructs identified within the CFIR framework to evaluate MaiKhanda's program theories, without necessarily having to develop a middle range theory. (However, a drawback in this case what that the choice of constructs was guided largely by the availability of data from MaiKhanda process evaluation). Secondly, I was looking at the intervention retrospectively rather than prospectively and developing a MRT retrospectively could be biased towards the observed results.

Theory based evaluation generally attempts to explain HOW the mechanism and context influences the outcomes of the intervention under study. Here there is an assumption that the intervention has worked or has had a positive effect. In cases such as the MaiKhanda intervention, the study did not yield conclusive outcomes for newborn mortality, which could help hypothesize the

'mechanism' of the intervention. In the absence of any discernible outcome, it is difficult to draw evidence-based theories as to how the intervention did not work.

Rather than trying to explain *how* the intervention did not work, I try to analyse *why* the intervention might not have worked by comparing it with well-informed hypotheses that outline the essential elements of successful QI interventions. I look at existing literature on the mechanism related to successful QI interventions. I identify key intervention strategies from other programmes, which generated these mechanisms and compare it with MaiKhanda data to identify key programme strategies that might be common across the various interventions thus making the programme mechanisms comparable across different settings. This process will help refine the original programme theory. Interventions are innovative but intervention theories are repetitive and boring. Program theories are comparable and the purpose of this comparison is to build the body of knowledge around the intervention(202).

Peter Pronovost and colleagues developed a successful QI Collaborative in ICU facilities in the State of Michigan in US, that successfully brought down central venous catheter (CVC) blood stream infection (BSI) rates by about 66% and was sustained throughout its 18 month study period(173). Mary Dixon-Woods and colleagues developed an ex-post theory of this successful Michigan study (known as the Keystone Project) (67). They developed well-defined hypothesis of what might have contributed to the success of this QI intervention. Their paper hypothesizes possible mechanisms of how the interventions at Michigan might have worked. I use ex-post theory developed by Dixon-Woods and colleagues, to explain the mechanisms of the QI intervention at MaiKhanda.

Using the evidence synthesized in Step 3, I will organize and compare the key strategies from the MaiKhanda study with the key ingredients(strategies) associated with the 'mechanisms' identified in the Michigan study. For instance, trying to compare if MaiKhanda intervention was able (or not) to develop a strong networked community, as observed in the Keystone Project

in Michigan, despite having similar Collaborative learning sessions. As MaiKhanda intervention strategies did not result in effectively 'triggering' the mechanisms, I will take a step deeper to analyse the implementation theory and trying to explain if the dose, duration, intensity and specificity of the intervention as well as the implementation fidelity was sufficient to 'trigger' the expected mechanisms. In doing this, I explore the relation between an interventions programme theory and its implementation theory.

4.3.5 STEP 5 Evaluating the implementation theory

Hulscher and colleagues (100) identified the key structural elements such as the selection of collaborative topics, the experts involved, the composition of the collaborative groups, the improvement steps undertaken and the type of activities supported that are determinants of success for QI Collaboratives. Embedding these intervention characteristics within health systems is key for the scaling up and sustainability of these programmes. QI Collaboratives are complex interventions being implemented within a complex adaptive health system. The complexity of the intervention implemented at MaiKhanda could potentially influence the outcome but this was not captured through the impact evaluation. In this final stage of the research strategy, I analyse the implementation theory within a complexity framework adopted from the work of Pawson (202). The framework categorizes complexity in terms of Volition, Implementation, Context, Time, Outcome, Rivalry and Emergence (VICTORE). I begin each section with a brief description of the components followed by an analysis in terms of MaiKhanda's implementation theory.

By complementing impact evaluation with theory-based evaluation, I hope to develop an insight into the context and mechanism by which the intervention was delivered (or failed to deliver).

Chapter 5 Evolution of MaiKhanda's Theory of Change

This section covers the theory of change linked to quality improvement interventions within the MaiKhanda programme. There are different methods to draft the Theory of Change (also sometimes referred to as 'initial programme theory'). This can be done by conducting a workshop with stakeholders including the designers and implementers of the intervention, reviewing programme documentation and reviewing literature. Other approaches include researchers' own observation and inquiry (exploratory on-site research) during the different phases of the programme. MaiKhanda 's initial Theory of Change discussed here is based primarily on review of the programme documents namely the project proposal at the inception of the project (Table 12-reference I) and the strategic review that was conducted towards the middle of the programme (Table 12-Reference XVII).

The aim of the MaiKhanda programme was to measurably reduce maternal and neonatal morbidity and mortality in Malawi over a period of five years. The programme intended to achieve these objectives by implementing programmes at both the supply side (improve service provision) and demand side (community participation) level. On the supply side the programme would improve the quality of skilled attendance and Emergency Obstetric and Newborn Care services, by working on maternal and newborn change packages, conducting deaths audits and data improvement. On the demand side, the programme would improve community preventive and care seeking behaviours by mobilising communities through women's groups. The plan was to implement five core activities in three districts – Lilongwe, Kasungu and Salima. These core activities included:

- Quality improvement in health facilities,
- Data improvement
- Structured support for conducting death audits
- Facilitation of women's groups at community level
- Population level surveillance for impact evaluation.

For the purposes of my PhD, I am only considering the supply side of the intervention (ie Quality Improvement in the health facilities) for newborn care. I have provided rationale for this in Section §1.10. Next section covers MaiKhanda's Theory of Change for quality improvement Collaboratives in the health facilities.

5.1 The original MaiKhanda Theory of Change (2006-2008)

MaiKhanda's Theory of Change is based on the premise that more than two-thirds of the delay in the utilization and provision of maternal and newborn care happens at the health facility due to a range of complex health systems related factors. Health service failures could include incorrect treatment, poor staff attitude, delays in referral process, poor co-ordination between various types of healthcare providers resulting poor quality of care. Shortage of staff and heavy workload can also overwhelm and discourage individual health care provider. Another constraining factor is the health care provider knowledge and skills. Also, the transient nature of the health workforce in Malawi creates severe weaknesses in institutionalizing knowledge within health facilities(208).

The interventions were to complement Malawi's 'National Road Map' for reducing maternal and infant mortality, by improving the coverage and quality of basic and comprehensive emergency obstetric and newborn care services.

MaiKhanda's Theory of Change included a hybrid quality improvement model that laid a great deal of emphasis on the choice of improvement methods to be used in the intervention. This integrated model for improvement in its operational version consisted of criterion based audits and standards of care integrated with the rapid improvement cycles (commonly referred to as PDSA cycles) (Figure 4)

5.1.1 Rapid Cycle Improvement

By giving the opportunity to front line health care providers to find local solutions to their problems, it is assumed that they will be major participants in the design and improvement of the care processes for newborns. This approach provided a structured learning environment at the point of care, to identify obstacles to access and delivery of care, and to document innovations that have proven to be effective in improving care. Traditionally it has been built around providers' knowledge base and use of guidelines is mainly through top down information from technical experts from the MoH or from MaiKhanda..

The operationalization of IHI's model for improvement was a composite intervention consisting of multiple components such as identifying the drivers of change, identifying the change package, forming QI teams, collaborative learning sessions and action periods with rapid improvement cycles (PDSAs) and run charts.

5.1.1.1 *Drivers of Change*

The intervention consisted of drivers of change ie the key set of strategies that were deemed necessary to bring about changes in maternal and newborn mortality. The project had identified four key drivers of change. These were:

- Clinically excellent care for mothers and neonates
- Women friendly care
- Information driven decision making
- Effective support systems

The bottom 3 drivers of change were common across for mother and newborns while there were specific secondary drivers for clinically excellent care for newborns.

5.1.1.2 *Change package*

The change package is essentially a list of ideas to stimulate the primary or secondary drivers of change to be tested in the facilities. The change packages were developed by the health facility QI teams and were closely

linked with the standards of care supported by CBA. The standards of care are broken down by the package into actions that the providers can do. The change package for clinically excellent care for newborn included Kangaroo Mother Care for low birth weight, neonatal resuscitation for asphyxiating babies and clinical management of newborn sepsis.

Examples of ideas included in the change package were: separating sepsis patients from other patients, educating patients on the importance of washing hands and applying soap. At the beginning of the QI implementation it was reported that facilities chose areas to work on in the change package on the basis of 'gut feeling'. However, gradually ideas to test became more 'data driven' from their own facilities and also based on evidence from literature.

5.1.1.3 QI Teams

Quality Improvement (QI) teams were responsible for implementation of change packages in the health facilities. They were the front-runners for testing innovative ideas in the facilities and formed an integral part of the QI implementation programme. The idea behind setting up QI teams was to harness local knowledge and build local capacity to solve their own problems.

The teams were multi-disciplinary and included representatives from the various providers within the facilities, for example, QI team members typically included a Matron, Nurse-Midwife, Clinical Officer, Lab Technician, and Data Clerk. However, a few of the QI teams, such as the team in Kasungu were developed from the existing MDR (Maternal Death Review) committee.

Each QI team had a leader who would regularly liaise with MaiKhanda to inform them of the QI work, submit data, prepare for workshops and arrange meetings. QI teams were not established in all the facilities given the staff turnover in the health centres as well as being rotated within the larger health facilities.

There were a number of ways that new people could get recruited to the QI teams, these included people expressing an interest in joining or being drawn

into the QI work to help solve a particular problem and then deciding to become a permanent QI team member.

5.1.1.4 QI Collaborative workshops

The workshops are also referred to as Collaborative workshops or "Breakthrough Collaborative Series"(209). During these workshops the facilities were taught about QI theory, and received training in the QI tools, such as PDSA, Maternal Death Reviews (MDRs), Criterion Based Audit (CBA) and data improvement. The initial QI implementation work began with QI team members from the 9 CEmOC facilities attending MaiKhanda workshops. Later on Health Centre Collaboratives were established in all the 3 districts.

Typically the QI workshop lasted two and a half days. It involved a 'Storyboard Session', which consisted of a board depicting the facilities QI work in the form of progress graphs, presentations of QI tools they had developed, and their aim statements.

A wider audience such as the Medical Director, representatives from MoH, and CHAM were invited to observe the Storyboards that allow them to learn about QI work going on in the facilities. This was important for improving the QI knowledge of stakeholder so as to have a wider engagement with them to get the critical buy-in during the scaling up phase later on.

The workshops also gave the facilities opportunities to meet with organizations such as the Malawian Blood Transfusion Service, who were potential stakeholders in resolving some of their resource issues. The Collaboratives also gave the QI teams the opportunity to meet with each other, share ideas and discuss problems and identify common solutions to their problems. The facilities were also able to learn about other facilities QI work, and adapt their ideas if appropriate.

The collaborative workshops served a dual objective of engaging the QI teams in improving the process of care by motivating staff to present their ideas and share ideas with other QI teams. It also engaged them in the content of care by providing hands-on training such as life-saving skills.

5.1.1.5 *PDSA Cycle*

PDSAs provide QI teams with a practical method and tool to materialize their ideas into action. The ideas from the change packages are tested using PDSA cycles. PDSA Cycle stands for Plan-Do-Study-Act. This includes 'Plan' to answer a question and 'Plan' for collection of data to answer the question. 'Do' refers to the carrying out of the change or test. 'Study' involves complete analysis of data. And 'Act' refers to reflecting on what changes are to be made and then plan for the next cycle. The PDSA is run to test different ideas generated by the facilities. .

5.1.1.6 Run Chart

Once the PDSA had been run, the facilities have to make sure that there is a certain level of reliability before that idea can be adopted as an improvement. Reliability is usually measured by the use of Statistical Process Control (SPC) charts also referred to as 'run charts'(190). QI offers a quantitative approach to review intervention mechanisms thru the use of these run charts.

5.1.2 Criterion-based audit (CBA)

The approach was led by LSTM (Liverpool School of Tropical Medicine) and it involved five steps: agreeing on standards of care, measuring current practice, identifying problems encountered, introducing change, and repeating measurement of practice. The process in facilities began by coaching staff to review existing relevant guidelines. Once the review team "owned" the standards of practice they would be coached in using their local data and their own knowledge of the situation to identify local solutions to their problems and repeat the cycle as often as needed till the level of desired standards were achieved.

5.1.3 The integrated model for improvement

The Criterion Based Audits and Rapid Cycle Improvement were acknowledged to be highly complementary of each other and it was decided to implement them through an integrated model in such a way that the process would begin with an audit cycle and would then move into rapid cycle change, once problems had been identified (Figure 4). The approach was to

be ‘bottom–up,’ participatory, and linked with MoH recommendations for finding solutions to the issues identified through death audits.

The reasons for integrating these different models were not clearly documented, especially since both LSTM and IHI had different perspectives on what constituted improvement and how it could be achieved. The Criterion Based Audits approach was based on quality assurance principles and aimed towards attaining minimum standards of care while Rapid Cycle Improvement was based on quality improvement principle and focused on continuous improvement. Further refinement of the model was suggested to ensure full integration. This was to be done through a small joint workshop between partners before the start of the interventions. However in practice, the CBA supported by LSTM and the improvement model supported by IHI were implemented as separate interventions particularly in the baseline and pre-intervention period.

5.1.4 Data Improvement

Data improvement is quite central to QI interventions. Improving data quality gave managers and clinicians the ability to make improvements on a continuous basis by accurately analysing events as they occur. Work on data improvement was to begin at the most basic level in facilities so as to understand health care providers needs and conditions around collection and immediate display of critical information. Work to improve the quality of information was to be conducted throughout the entire period covering all of the three districts. It was assumed that interventions at the facility level would build capacity in individuals for analysis, improvement and monitoring of basic processes of care. Experience from these initiatives would then be shared with all District Health Officers and MoH leadership through an on-going data improvement collaborative during quarterly meetings at the DHO. This would lead to changes in procedures and practices, which would establish permanent improvement for the facility despite turnover of staff. The interventions were to be implemented across the full vertical slice of the health care system. A vertical slice of the health systems included the randomly

selected health centres, rural hospitals, District Hospital and the referral (CEmOC) facilities.

5.1.5 Support for death audits

Death audit was a separate unit of activity using the standardized forms provided by MoH. The intervention would support and enhance the reporting of audit results and encourage the use of data from audits in the improvement process as an integral part of the facility coaching. There was a greater emphasis on maternal death audits than newborn death audits.

The process of a maternal death review (MDR) typically involved analysing one maternal death at a time followed by a discussion focusing mainly on the 'action points' or activities that need to be followed up from the review. The review team would then target specific ideas that the facilities could take to prevent such a death from recurring in future. Although challenging, an important part of MaiKhanda programme was to introduce the culture of no blame during MDRs.

MaiKhanda played a facilitatory role in conducting the death audits. The district teams conducted the death audits. In the pre-intervention period, MaiKhanda played an active role in providing technical support to the review process through LSTM clinical team. The project adopted a more passive role in the early and late intervention period focusing more on engaging facility improvement teams to take action on the issues arising out of the death audits, using PDSA techniques.

Although facilities struggled to conduct death audits on a regular basis, this was a well-recognized activity institutionalized within the Malawian health care system as this was a direct recommendation from the MoH for district facilities.

In simple terms, death reviews provided a reactive approach for improvement, i.e. action was initiated once an event has occurred, while criterion-based audit and Model for Improvement adopted a more pro-active approach towards improvement.

5.2 Implementation

Here I provide a description of how the intervention was to be implemented on the ground. An examination of the original programme proposal and documents from the strategic review conducted by the end of phase I (2008), provides an understanding of the key characteristics of the intervention as well as organization of work associated with it. This provides the basis for formulating the implementation theory. This section covers key areas such as planning and organization of work, stakeholder engagement, district selection, partner role, governance and oversight, sustainability, communication and risk management in the project.

The intervention was to maintain close links with the Ministry of Health (MoH) through the Reproductive Health Unit (RHU) and the Safe Motherhood Task Force (SMTF). The assumption was that this would accelerate sharing of programme results, both successes and failures, to benefit others at work on this problem.

5.2.1 Planning and organization of work

The improvement team was began work initially in CEmOC facilities as a way of learning about the system and providing support to the clinical leadership in these facilities. Once the districts were selected, preliminary work began with District Health Management Teams (DHMTs) in order to familiarize them with MaiKhanda's approach to quality improvement and learn from them what problems they wish to address first. The exact design of district work was to be determined by the DHMTs and by advice received through local advisors and through stakeholders meetings.

Beyond the District Hospital, the phasing of facility work within each district, involved two-thirds of the district in the first phase and included both public and CHAM[xi] facilities. While the intervention was intended to spread more organically to the newer facilities, in reality, this was influenced largely by the

[xi] *CHAM (Christian Health Association of Malawi) facilities have a service level agreement with MoH and are responsible for approximately 30% of the health service provision in Malawi.*

randomization principles of the evaluation design. The plan was to work across a 'vertical slice' of the facilities. Except the central level CEmOC facilities, all the rest of the facilities were classified as BEmOC facilities (although they were not fully functioning as one) and was subjected to a randomization process, so as to be assigned as an intervention or 'control' facility.

Representatives from the intervention facilities were invited to a QI Collaborative learning session where they were introduced to QI concepts through 'storyboard' sessions and initiate a process of sharing ideas and improvement across facilities. This was followed by support to local QI teams through on-site coaching by MaiKhanda QI officers.

The initial plan was to coach one district at a time, with very frequent visits taking place over the first 2-3 months, decreasing after that time. Collaborative learning sessions mainly with support from Consortium technical advisors were to be held every quarter, followed by intensive coaching of district one after the other. This meant that initiation of improvement activities would be staggered across districts. However, in practice, all the facilities selected through the randomization process were initiated into the Collaborative learning sessions together. Thus there was a CEmOC Collaborative learning session and a district Collaborative learning session for each of the 3 districts held simultaneously. These session were held on a quarterly basis (known as 90-day cycles). Throughout the implementation period, there were challenges with following this plan, for a variety of contextual reasons, discussed later in the Analysis Chapter (§7.2.3). This influenced the implementation strength of the interventions and consequently trial results as well.

5.2.2 Stakeholder engagement

Important stakeholders were engaged in completing the design of the QI intervention. The stakeholders were oriented to fully understand the model being developed for quality improvement of maternal and neonatal health (MNH) services and would be kept aware of programme developments.

The primary stakeholders identified for the QI intervention were the health care providers from the facilities. The aim was to train and support local leaders so that the principles of improvement became part of the culture of the local health system. The collaborative model would ensure that 'local champions' developed their skills as they taught new entrants into the collaborative. Modest continued support, in the form of funds to attend meetings and continued learning and coaching in statistical techniques would ensure the continuation of the activity over time.

District Health Management Team (DHMT) would be involved in the original planning and coordination of programme implementation so that initiatives developed from QI collaborative and audit cycles could be incorporated into DHMT plans. The programme also planned to develop leadership skills within the DHMT.

Other stakeholders involved Zonal health officers and Central level MoH officials, in particular the Reproductive Health Unit (RHU), the Nursing Department and other relevant agencies such as the Health Information System unit.

5.2.3 District selection

District selection was done in discussion with the Ministry of Health (MoH). The Consortium was asked by the RHU to work in Districts that are among the "ten worst" based on reported maternal deaths and case fatality. There were practical issues and political sensitivities associated with district selection. Department of Nursing wanted to consider Salima and Ntcheu districts. The Consortium finally started work in Salima, Kasungu and Lilongwe. There is no documentation on the rationale for the selection of these districts. Nevertheless, this was discussed and reviewed with MoH and finalized during the initiation of the programme in February 2006. The choice of districts and the process that followed had implications for the 'ownership' of the programme, which will be discussed in the Analysis Chapter (§7.1.1). An important point to note here is that Lilongwe district is larger than Salima and Kasungu district put together and has a sizable urban population the size

of a district. However, the choice of districts was beyond MaiKhanda's influence and consequently having Lilongwe as one of the intervention districts affected the programme planning and budgeting and implementation.

5.2.4 Partner role

LSTM (Liverpool School of Tropical Medicine) provided expertise in criterion based audits and deaths audits to health facility staff. IHI (Institute for Healthcare Improvement) had expertise in quality improvement in health care and were instrumental in introducing QI techniques such as the rapid cycle improvement in the selected health facilities. Their experience of quality improvement in healthcare was mainly from the developed countries such as US and Europe.

LATH (Liverpool Associates in Tropical Health) a subsidiary of LSTM and registered as a local entity in Malawi, provided the local team with logistical, administrative and financial management support.

A small office was maintained in Lilongwe to accommodate technical and administrative staff, with district offices being established later on in the project. The assumption was that the Collaboratives would trigger a diffusion of the improvement model through the local health facility leadership that was engaged during the Collaborative process. In reality, the spread of the intervention was not as organic as it was anticipated and district offices[xii] had to be extended to build local capacity and momentum for quality improvement.

5.2.5 Governance & oversight

The programme was managed in Malawi through LATH who were responsible for local management of the Consortium[xiii]. The team in Malawi was headed by a programme manager who was supported by a programme officer dedicated for QI activities. The team received technical support from

---

[xii] *The District Offices for MaiKhanda was established earlier on in the project but this was used by the community intervention and the M&E team. QI team operated centrally from Lilongwe office.*

[xiii] *MaiKhanda had not registered as an NGO by then.*

Consortium members. The programme implementation team at MaiKhanda[xiv] was responsible for day-to-day activities of the Consortium and also for maintaining the active relationship with the MoH.

A Programme Steering Committee was to provide hands-on oversight of local operations. This committee would include Consortium representatives and Malawian advisors active in championing the development of the safe motherhood and neonatal health improvement, e.g. obstetricians, midwives. They would support staff in dealing with anticipated problems such as slow implementation, resistance by local facility staff, personnel problems. They would advise on the continued interaction with the Road Map and with other stakeholders interested in this area as well as on major decisions such as solicitation of additional funds. In terms of programme implementation, this Steering Committee never got established.

A Consortium Management Committee consisting of a member each from the Consortium partners and The Health Foundation met at least twice annually and reviewed major decisions such as changes in the scope of work or other strategic decisions such as seeking additional funds, extension of the programme of work. This later on was renamed as the Programme Management Board and included along with representatives of the consortium partner and donor, the senior management team from MaiKhanda.

5.2.6 Sustainability

The intervention planned for a sustainable model of information sharing and continued learning of improvement and adaptations through the establishment of 'knowledge agents'. The knowledge agents would gather, aggregate and analyse data and act as the connective tissue for the programme, driving knowledge exchange and interpersonal connections and ensuring continuous learning. These agents would use mobile technology (cellular phones, handheld computers) to collect data and knowledge, travelling from site to

---

[xiv] *The term MaiKhanda did not come in to effect till the organization was registered as a NGO in 2008. Previously it was referred to as ThfC-The Health Foundation Consortium.*

site, providing caregivers with locally useful data and proactively carry to them resources and answers to pressing questions. The plan was to locally recruit from the existing cadre of front-line health facility staff so as to ensure programme sustainability. At an operational level, this was done by the Facility Intervention officers (improvement advisors) from MaiKhanda during their field visits to the health facilities during the 'Action Period'. From a sustainability perspective, this was a challenge as the MaiKhanda officers were involved with, but not part of the health system in which the intervention was being implemented.

5.2.7 Communication

The Collaborative learning sessions, it was assumed, would allow for effective communication at different levels of the health service. The programme envisaged exchange of data, progress reports and ideas among care providers, experts and actors in the care delivery system, through the collaborative learning sessions. Using a common set of measures to track progress, participants could significantly reduce cycle times for local implementation, generating a growing list of improvements and adaptations.

At the programmatic level, programme framework would include objectively verifiable indicators to monitor progress in relation to programme inputs, process and outputs. Partners had access to the indicator reports and thus would be able to use the indicators to continually improve the programme through their participation at the programme oversight level. However, this process of continuous improvement sometimes required considering an intervention redesign, which was in conflict with the evaluation design principles of maintaining standardized interventions.

5.2.8 Risk and risk management

A log frame was developed as part of the planning process which addressed the risks and assumptions at each level of activity. The Consortium reviewed and discussed potential risks to the programme as part of the planning process. They identified both the external and internal risks.

External risks

Political instability

There was the risk of political instability at the start of the programme, which could aggravate the shortage of resources at the local level. This would have implications for intervention implementation and subsequently the outcomes.

District support

Malawi was going through a process of decentralization at the time of implementation of the project. While this was well mapped out in theory, in practice it was not clear how decentralization would work. There was an assumption that the districts would be supportive of the QI work in their districts.

Human resources

The Consortium was well aware of the human resource crisis within the health sector in Malawi, at the start of the programme. However, substantial funding from DFID through the Emergency Human Resource package, much of it specifically focused on improving salaries, was expected to stabilise the situation(210).

Internal Risks

Three major internal risks were identified at the start of the programme and documented in in the project proposal:

- A partner decision to depart from the Consortium
- Disagreements arising as to one partner's performance
- Local disagreement with scale-up plan

The internal risks outlined provide an insight into the tensions that were prevailing between implementation partners even in the proposal development stage. These risks became obvious once programme implementation on the ground began. These internal risks affected the duration and intensity of the implementation process.

5.2.9 Additional Activities

A number of additional activities were suggested such as the leadership buddy system where each facility would be linked "buddy" to a high level leader from the health system (District Health Office, Ministry of Health, Reproductive Health Unit, and CHAM), exchange visits so that QI team could observe the improvement work going on in other facilities. However, these activities were not implemented during the project tenure.

There were other strategies such as promoting 'local experts' within facilities, who would then be able to teach their skills to other facilities and to new members of QI teams in their own facilities. The idea behind the strategy was that the facilities would respond better to learning from a peer as opposed to learning from outside experts. These local experts would be staff members with clinical skills who had become an expert in a certain area of QI. This strategy was first initiated in Nkhoma (CHAM) hospital and was successful within that facility but did not gather enough momentum or rather there was not enough capacity within MaiKhanda to replicate it beyond that facility. There was also the strategy of supporting 'super improvers'. The idea was that there would be 'champions' emerging from the work in the facilities-these were QI team members who were enthusiastic about improvement work. MaiKhanda would build capacity of these super improvers to conduct quality improvement (Collaboratives, mentoring, developing PDSAs). Their role would be to help expand QI work within their own hospital and within the health centres.

While these additional activities were well intentioned and very strategic, their execution remained sub-optimal. This was linked to MaiKhanda's own capacity but also shaped by the strategic vision of the partnership and the programme design.

The description above provides an overview of the proposed theory of change. The analysis will look at the changes to the theoretical and operational dimensions of programme implementation over the period of the study. The analysis will also look at how assumptions changed over a period

of time and subsequently affected programme implementation.

5.3 The revised Theory of Change (2008-2010)

It was evident from the baseline evaluation that there were challenges with the intervention design i.e integrated improvement model because of differences in partner views, which affected effective facilitation by MaiKhanda staff subsequently affecting programme implementation in the health facilities by the individual QI teams(33). The main point of divergence was LSTM's Standard of Care approach, which had its theoretical underpinnings in quality assurance (QA) whereas the rapid improvement cycles promoted by IHI had its underpinnings in improvement (QI) theory (Also see Box 1). MDR used a reactive approach to achieve standards of care while CBA was a more proactive approach. Nevertheless both models adhered to the principles of quality assurance. The fundamental difference between a QA & QI is that quality improvement focuses on continuous improvement of systems and processes used in an organization while Quality Assurance (QA) focuses on episodic survey of performance, addressing mainly the errors of individuals.

Coupled with this the donor affiliation was also inclined towards IHI Model for Improvement, having successfully funded IHI intervention in the NHS (UK). All this meant a revision to the theory of change, focused largely on the IHI model for improvement. Thus the revised model used in Malawi, was based on the IHI 'Model for Improvement' used in other developed country settings and emphasized on the following:

- Change packages and improvement cycles
- Breakthrough Collaborative Series (BTS) also referred to as 'Collaboratives".

  The main difference was in the approaches. While the original framework suggested death reviews or criterion-based audits as the starting point of the intervention, the revised model was a 4 step process:

1. *Aims:* Having clear aims for improvement so that local QI teams have a clear understanding of what they are trying to accomplish.

2. *Measures:* Measurement, especially local measurement that allows the actors to know when a change is an improvement (by creating a continuous feedback mechanism).
3. *Change Concepts:* Ideas for change to the systems that are alternatives to the *status quo* designs.
4. *Cycles of Local Testing and Adoption:* Opportunities for local organizations and teams to test those ideas for change in context so as to refine them and, if effective, stabilize them in place, or, if ineffective, reject them. These cycles are known as the PDSA (Plan-Do-Study-Act) cycles.

The assumption was that by adapting this MFI, lots of change ideas would be developed at the local level. Of these change ideas, some would bring about improvements, which had potential to be a local innovation, which would then be spread widely and more rapidly. This spread was intended to be done through the "Breakthrough Collaborative Series". The general architecture of a Breakthrough Series Collaborative (Figure 3) which includes QI teams attending in-person meetings ("Learning Sessions"); linkages through visits and electronic communications; mentoring of QI teams by "faculty[xv]" who knew the subject matter and improvement methods; systematic, transparent and shared local measurement; senior leadership engagement to remove obstacles to local QI teams' progress; peer-to-peer exchange, local adaptation, and highly participative management(211). A Learning Collaborative consisted of a series of 3 Learning Sessions over the course of 9 to 12 months in which organizations worked together towards a common aim. One of the cornerstone philosophies of a Learning Collaborative was "everyone teaches, everyone learns"—promoting the spread of successful ideas from one QI team to another. In the original framework, the role of Collaboratives was mainly focused on orientation and training on the QI methods.

[xv] *This would include the local champions from the facility but it was mainly MaiKhanda staff who were trained in IHI's improvement methods.*

By bringing 'communities of peer groups' together in the Collaborative sessions, for shared learning, the Collaboratives tried to create an environment where people were comfortable sharing their local innovation and improvements using data from their facilities. Equally important was what happened in between Learning Sessions; these were called Action Periods, when QI teams tested and implemented new ideas as well as collected and submitted monthly data and reports. MaiKhanda officers provided technical support to these communities of people (QI teams) in each facility during action periods, to harness their change ideas into improvements. By continuing this iterative cycle over a period of time, the programme expected to develop a critical mass for improvement, which would then support broader systems level change.

Indeed, by September 2008 the donor agency decided to follow the model for improvement promoted by IHI, leading to the exit of LSTM and LATH from the Consortium. In operational terms, this meant the departure of the integrated improvement model which included death reviews and criterion based audits as intervention starting points and a greater emphasis on development of the change packages, improvement cycle.

Some of the major changes to the design, as a result of this revised theory included a one improvement methodology i.e the use of the *Model for Improvement(46) and* data driven decision making for technical work as well as for programme management. There was also a change in the governance structure with a greater emphasis on strong in-country leadership through the appointment of a Programme Director for MaiKhanda. The programme Director main roles was to connect the technical leadership of advisors with the management and work of the local technical team and to lead MaiKhanda in advocacy with MoH and other stakeholders to ensure that the programme influenced and adapted to the dynamic policy environment.

The composite set of interventions from IHI's improvement methods had been well-tested in the United States, Europe, as well as in some resource constrained settings such as South Africa and Peru. The key activities

included a customized collaborative approach for QI with emphasis on solidifying the local QI team, intensive on-site coaching by MaiKhanda officers, standardized process for coaching visits to enable use of improvement tools and focussed meetings. Local skill building was to be done through on-site clinical coaching for QI and in-situ training for clinical skills. The revised theory of change also emphasised on leadership development by enrolling "super improvers" identified from within the facilities, to a training course. There was to be a greater focus on engaging DHMT as well as leaders from all levels of the system and incorporating QI as a supervisory tool. The data improvement activity was to continue as before but with greater coordination between the implementation and evaluation team.

5.4 Assumptions

Some of the major assumptions associated with the intervention was that:

- The Ministry of Health was eager to improve both the quality of data coming from health facilities and the use of data within those facilities to drive improvement.
- Staff at the facilities would be eager, enthusiastic and motivated to implement the Model for Improvement in their facilities.
- The subject matter knowledge and skills of providers in terms of maternal and newborn health were good.
- Through a structured series of small interventions, designed by the local participants themselves, the healthcare system could be improved from the ground up, and local needs could be accommodated, mostly through more efficient use of existing resources rather than demanding for additional resources.
- Staff attrition and resource availability were considered as challenges at proposal development stage. However, changes could be successfully implemented with minimal additional financial resources and technical assistance as long as there was strong local commitment, enthusiasm and leadership.
- There also seemed to be an implicit assumption that the health systems

building blocks[xvi] within the Malawian health systems were functional and that model for improvement would essentially improve the interaction between the building blocks.

- A lean management model, involving one QI officer to cover the 3 districts was proposed, assuming a rapid uptake of the intervention by the facility QI teams and subsequent organic spread of the model, given the level of engagement and communication with different layers of the health system.
- There would be a buy-in and rapid uptake of the model for improvement, so fast that it would contaminate the control sites identified for evaluation purposes, making it difficult to measure impact of the intervention.

5.5 Summary

This Chapter provides a description of the original theory of change as well as a summary of the proposed programme implementation followed by a description of the revised Theory of Change in the later part of the project which formed the mainstay of MaiKhanda intervention from 2008 onwards i.e the period corresponding with the period when the RCT evaluation design was in place.

The next Chapter synthesizes all the available evidence to provide a picture of the gradual evolution of MaiKhanda's Theory of Change and the factors that influenced it. This will also consider the changes to theoretical and operational dimensions of programme implementation over the period of the study. The following Analysis Chapter will also look at how assumptions were influenced by the context and intervention complexity over a period of time and subsequently affected programme implementation.

[xvi] *Based on WHO definition of health systems building blocks: Service delivery, health workforce, Information, Medical products, Vaccines and Technology, Financing, Leadership and Governance.*

Chapter 6 Evidence Synthesis

6.1 Collating and summarizing the evidence

The evidence synthesis is preceded by Step-2 (§4.3.2) in the research strategy and includes the collation and summary of the individual evidence. This process of collating and summarizing the evidence consists of two main parts:

The first part covers a descriptive summary of all the process evaluation studies that were conducted as part of MaiKhanda evaluation to understand the 'black-box' of the intervention. The second part is a summary of the information available through various project documents such as the original proposal, the monitoring reports, the field visit reports, the bi-annual progress reports, the programme management board meeting minutes as well as documents such as the 'data deep dive', which is essentially a reflective evaluation conducted jointly by the evaluation and implementing team to collate and review and reach a consensus on what actually happened with the intervention.

6.1.1 Process evaluation studies

The evaluation approach for my PhD used the process evaluation data from the MaiKhanda study (§1.8) to outline the specific activities within the QI intervention, analyse their processes and mechanisms and also took into consideration the context within which the interventions were implemented. Table 10 provides an overview of the QI intervention components and their data sources used for the theory-based evaluation

6.1.1.1 Process

The process for the QI intervention consisted of collaborative learning session and action periods. There were also separate data improvement workshops and maternal death audits being conducted in the facilities.

The dosage and implementation of the intervention was evaluated using the routine monitoring data from the intervention team. However, this data was

either incomplete or unavailable and subsequently data reported during the Data Deep Dive provided some additional information on the intervention process.

	Quality Improvement Intervention	Data source
Process	Learning Collaboratives & action periods	Project monitoring data
	Data improvement	Monitoring data, Data workshop reports
	Project Implementation	Quarterly Progress Reports
	Implementation of change packages	Implementation Monitoring data
Mechanism and Context	Composition of QI teams	Staff psychology survey CEmOC survey Provider Knowledge survey
	Resource availability	Health Facility resources survey Health facility surveillance

Table 10: Secondary data sources from MaiKhanda process evaluation

6.1.1.2 Mechanism and Context

One aspect of the QI work was to motivate and encourage staff to work better in the face of the adversities they faced on a daily basis (lack of supplies, increased workload). In order to understand the attitude and outlook of health care providers, a survey of members of staff at all 78 health facilities (including facilities in the comparison areas) in the three districts and 15 in-depth interviews were conducted between January and May 2008.

A survey of all the non-randomized health facilities (CEMOCs) were also conducted at the beginning of the study to assess the QI activities being carried out in these facilities and to gauge the understanding and general impression regarding model for improvement among QI team members.

It has long been recognized that skilled attendance at delivery can prevent maternal and neonatal deaths(212). The process evaluation included a

provider knowledge survey, designed to assess providers' knowledge of best practice, as well as to explore providers' perceptions of the care they provide, their confidence with key subject areas and the effect of training on their knowledge. The questionnaire was designed to identify the degree to which providers' knowledge; confidence and training could explain observed case fatality rates(213).

The quality of service provision at health facilities is dependent on the context within which the services are provided. One of the main contextual factors determining the quality of service provision, in the Malawian health care setting was the availability of human and material resources. The availability of material resources included the availability of physical infrastructure and the availability of drugs and supplies. The availability of essential drugs, supplies and BEmOC and CEmOC signal functions was assessed on a monthly basis using the Health Facility surveillance tool (33). The availability of human and other material resources such as the physical infrastructure was determined by the Health Facility Resources Survey (33).Measures such as the nurse to population ratio; provision of water, electricity and referral services, were looked at.

6.1.2 MaiKhanda programme documentation

The evaluation for the MaiKhanda interventions was conducted independently by the team from UCL. A reflective evaluation of the quality improvement interventions towards the end of the project was done to gather a collective knowledge of the impact of MaiKhanda's interventions. This reflective evaluation included the implementing partners-IHI & WCF, the evaluation partner-UCL and the donor-The Health Foundation playing a mediatory role. The process was known as the 'Data Deep Dive' exercise and two such consultative meetings were held-the first such consultative meeting was conducted in July, 2010 and the subsequent meeting was held in October, 2010. A summary of the key observations, reflections and findings from the meeting minutes and Data deep dive documents was collated and analysed to provide insight about the perception of the implementing partners as well as

contextual and other factors that are likely to have influenced the interventions on the ground.

Other key documents included the original proposal document, the revised strategy document, the biannual progress report and minutes of the Programme Management Board meeting from 2007 until 2010, field visit notes from Technical Advisors.

Table 11 provides a list of all the documents used in evidence synthesis. The column on the right of the table provides a numbering that has been used for referencing in this Chapter.

Table 11: List of source documents used in evidence synthesis

List of documents used in the study	Reference
I. Original Proposal Document Oct, 2006	I
2. Baseline Survey June, 2006	II
3. Baseline Evaluation Aug, 2007	III
4. Health Facility Resources Survey 2010	IV
5. Provider Knowledge Survey 2009	V
6. Staff Psychology Survey 2008	VI
7. CEmOC Survey 2008	VII
8. Women Friendly Care Survey 2010	VIII
9. Data Deep Dive (Reflective Evaluation) Aug, 2010	IX
10. PMB Minutes Sept, 2006	X
11. PMB Minutes Feb, 2007	XI
12. Evaluation Report June, 2007	XII
13. PMB Minutes Oct, 2007	XIII
14. Bi-annual Progress Report Dec, 2007	XIV
15. PMB Minutes Feb, 2008 and Apr, 2008	XV
16. UCL Evaluation Response	XVI
17. Bi-annual progress report (Including Strategy review document), June 2008	XVII
18. Handover notes to MaiKhanda Programme Director	XVIII
19. PMB Minutes Oct, 2008	XIX
20. Director's Report Feb, 2009	XX
21. PMB Minutes Feb, 2009	XXI
22. Quarterly Progress report May, 2009	XXII
23 QPR July,2009	XXIII
24 PMB Minutes Sept, 2009	XXIV
25 Quarterly Progress report Oct, 2009	XXV

26 Quarterly Progress report Jan, 2010	XXVI
27 Lloyd Provost Field Visit Feb, 2010	XXVII
28 Quarterly Progress report Apr, 2010	XXVIII
29 Quarterly Progress report July (Including St. Gabriel Case Study), 2010	XXIX
30 Strategic Review Meeting notes Sept, 2010	XXX
31 PMB Minutes Mar, 2010	XXXI
32 Partners Meeting Oct, 2010	XXXII

6.2 Consolidating and Synthesizing the evidence

All this data was collated and summarized into an excel spread sheet. The data thus generated from the various sub-studies as well as the various documentary information available throughout the course of the project, were organized according to the domains and constructs from the intervention framework (Figure 20). The five major domains covered include the intervention characteristics, health systems context; MaiKhanda's own internal settings, the individual characteristics of the (human) agency involved in QI and finally programme implementation. By organizing the data in this manner, it provides a comprehensive picture of MaiKhanda QI programme action using a mix of qualitative and quantitative data. This is followed by a synthesis of all the available data within each of the constructs, identified within each domain, according to different phases of the intervention. There were four distinct phases to MaiKhanda intervention: the baseline, pre-intervention, early intervention and late intervention period. The reasons for splitting the intervention period into early and late phases have been described in detail in Chapter 1 (§1.6). The synthesis document is presented in Table 12

The next section is a description of each of the synthesized constructs.

Table 12: Evidence Synthesis using Consolidated Framework for Implementation Research (CFIR)

	Intervention Periods	Baseline period (June 2006-January, 2007)		Pre-intervention period (February, 2007-September 2008)			
	Years	2006		2007			
	Quarters	Q3	Q4	Q1	Q2	Q3	Q4
A	**Intervention characteristics**						
1	Conceptual clarity	MaiKhanda improvement model integrating the Criterion-Based Audit (or standard-based audit) cycle and the Rapid Cycle process improvement (or PDSA). The rapid cycle process improvement is based on IHI Model for Imrpovement which has its theoretical underpinnings in Deming's system of profound knowledge.[1] It was felt that an RCT design for evaluation would be contaminated by the rapid spread of the intervention to the control sites.		It was decided that the programme would not fund procurement of drugs and medical supplies, but would support the QI activities already initiated to ensure systems for drug and supply inventory and ordering were developed at district and health facility level[11]. Essential support systems was one of the 4 proposed change packages. This was because of MaiKhanda's position as a 'facilitator' rather than 'implementer'.		The baseline evalaution reported confusion among Malawian staff about the integrated model, particularly the links and overlap between criterion-based audit (CBA) part of the LSTM-led maternal death reviews, and Plan-Do-Study-Act (PDSA) part of the IHI-led QI initiatives[3,12]. There were uncertainties expressed with regard to the effectiveness of the combined QI Model. The senior technical team felt that the integrated model did not work in practice leading to confusion among MaiKhanda team members on the ground.This confusion extended to facility based teams as well. Email from THF Director: *"In reality we have seen that the work in Malawi has been frustrated by the mixed approach"* [14].	
2	Intervention design	The integrated model of criterion based audits and rapid PDSA cycles Death Reviews and Data Improvement form the basic components of the QI implementation programme at the health facility level[1]. The programme would be delivered across health centres and CEmOC facilities in 3 districts over a period of 5 years in a phased manner.	While the intervention(s) were designed within the framework of a RCT, the 9 referral CEmOC facilities was not meant to be part of the randomization (control) process. It was decided to implement the intervention across the 9 facilities as a health system strengthening exercise (for ethical and methodological reasons)[12] MoH recommendations led to a national pattern of implementation of the Maternal and Neonatal Death Review (Audit) programme. MaiKhanda programme would complement this MoH strategy by facilitating the local staff to conduct audits in a blame-free and participatory manner. The information generated by the review was to be used to determine the focus of the criterion-based and rapid-cycle improvement activities. Although MDRs were taking place in four of the nine facilities prior to the work of MaiKhanda, the QI intervention helped to strengthen the existing MDR structure[12]. Newborn health was largely absent from early QI discussions.An increase in newborn deaths has been observed, probably through improved reporting. Newborn death audits were not routinely done in the facilitiesnor was it prioritized in the MaiKhanda interventions [12]. This model is used as a tool for improvement of quality of care in a 'bottom-up', participatory approach and is in line with MoH recommendations. It was realized that it might not be possible to define and develop QI interventions at the very beginning of the project given the wide variation in context in different intervention settings. An initial exploratory phase was therefore suggested in determining the factors that need to be focused on in a given context. Hence the pre-intervention period [12]			Currently, the Programme has 37 staff members but only 1 QI officer. The programme is in the process of recruiting new QI staff members (that was not in the orignial proposal): Data Coordinator , Assistant Programme Officer (Quality Improvement) and 3 additional District Quality Improvement Officers[15].
3	Embeddedness	Leadership buddy system, developing local leadership, QI champions, exchange visits were activities planned in the original proposal for penetrating QI into the health system[1].	One of the CEmOC facilities reported encountering resistance for their QI activities from another organisation working in their facility. They explain "that their newly developed labour graph met resistance in the sense that the Safe Motherhood Initiative authorities said we should hold using it"[12].	MaiKhanda's interventions aims to work alongside the Malawian Ministry of Health (MoH), in achieving their national 'Road Map' for reducing maternal and infant mortality. This includes - developing & implementing appropriate behaviour change interventions, conduct maternal death reviews and clinical audit & strengthen health facilities at all levels with functional EmOC services [12]. The collaboratives and leaders appointed within the nine CEmOC facilities are expanding the intervention across the outreach health centres. Maternal death reviews in health facilities, in line with Ministry of Health needs have been used to identify problems and implement solutions[12].			Strategy is to work closely with both health centres and DHMT so that critical needs can be incorporated and budgeted for in the district Implementation Plan[14].

	Intervention Periods	Pre-intervention period (February, 2007-September 2008)		
	Years	2008		
	Quarters	Q1	Q2	Q3
A	**Intervention characteristics**			
1	Conceptual clarity	IHI (CEO) suggested the Model for Improvement as a unified, core technical framework for managing change – The Model for Improvement was to provide the core framework for improvement activities undertaken by participating facilities and communities. General adherence to this approach was expected to limit confusion for program participants and staff as they seek to transform their outcomes.[14,15]. This change in programme concept was approved by the donor, THF.		IHI model for Improvement endorsed in the revised strategy phase[20].
2	Intervention design	A revised plan for Quality Improvement was proposed which consisted of the following components: Strengthened technical support to MaiKhanda Malawi and QI Teams. Quality Improvement (IHI Model) in health facilities. Clinical Skill Building. Data Improvement. Leadership Development [15]. The opting out of LSTM left a gap in the content side of the project such that activities related to criterion based audits and maternal and death reviews were some how affected. To address this problem, clinical advisors were identified and they were to provide mentoring and coaching of Quality Improvement teams on various issues[17].		
3	Embeddedness		MaiKhanda engaged in strategic MoH meetings such as SMTF (Safe Motherhood Task Force), QATF(Quality Assurance Task Force), DEC (District Executive Committee)[18].	

	Intervention Periods	Early intervention period (October 2008-September 2009)			
	Years	2008	2009		
	Quarters	Q4	Q1	Q2	Q3
A	**Intervention characteristics**				
1	Conceptual clarity		A revised plan was developed for the next phase of the project based on th erevised strategy. There was a directive from the MaiKhanda Director suggesting QI to be adapted as the ovararching programme strategy for all MaiKhanda interventions. This meant referring “quality” improvement (QI) as “facilities” intervention (FI)[20].	Consortium partners reaffirmed that the aim of the program is to reduce maternal and neonatal mortality, using quality improvement in the facilities, women’s groups in the communities, with synergy and coordination between them.Donor (THF) clarified that MaiKhanda is a decentralised program, and that IHI is the lead partner with a role in QI technical execution and other partners also have roles in executions of their technical areas. [22]	The revised plan emphasised leadership at district level and a decentralised approach[24].Salima district was proposed as an “Exemplary District” (ED.) FI Team believed that focusing their efforts on one district, would lead to a higher level of success in the ED which could then be documented and shared as proof of the MaiKhanda model’s success [21]. This was based on the "diffusion of innovation" theory.
2	Intervention design		Shift in programme strategy to reflect a QI approach throughout the program (community intervention, management support services, monitoring and evaluation. In operational terms it meant “getting women to give birth at the facilities and then ensuring that they receive a ‘skilled delivery’ once they get to the facility” as the primary objective of the project [20]	There was a prospoal to put more efforts into Salima district while maintaining the current level of support in Kasungu and Lilongwe[21].	
3	Embeddedness		RHU and DHOs involved in MaiKhanda's strategic planning meeting. Capacity stregnthening of CMED (MoH) on data collection, interpretation and dissemination[20].		

	Intervention Periods	Intensified intervention period (October 2009-December 2010)				
	Years	2009	2010			
	Quarters	Q4	Q1	Q2	Q3	Q4
A	**Intervention characteristics**					
1	Conceptual clarity				Difference about the primary objective of the intervention. Getting women to give birth at the facilities and then ensure "skilled delivery".However the donor clarified that this was not the primary objective[30].	Every project has a timeline – they need to start "alongside" and develop into "through." There is a need to prove the intervention during "alongside" and then spread with "through." It is not possible to implement, spread/scale up, or be sustainable without working through the MoH[32].
2	Intervention design		Because most maternal deaths occur in CEmOC's, the approach to randomization of health centers may not be able to detect the impact of QI. But as mentioned above, we may be able to work backwards from the CEmOC data to look at deaths by referring health center[27].	Salima: 31/36 planned HC QI visits (86%) achieved. Of the 31 visits, 7 visits:teaching visits on QI methodologies, 9: follow up meetings,16: coaching visits for QI teams in the HCs. During the quarter, 17/12 visits were made to Salima District Hospital. Of the 17 visits, 3 (18%) were MD audits, 6 (35%) for data collectionand 8 (47%) for coaching of the QI team[29]	QI interventions focussed largely on the CEMoC facilities and largely on mothers. (wheras based on evaluation, the emphasis had to be on the health centres and equaly among newborn and maternal)[9].	The project was originally designed to be scaled up to all areas of the three districts within five years. This scale-up phase was postponed because the partners felt the current interventions had not yet been successfully implemented[30].
3	Embeddedness	Many of the breakthroughs we had in "women friendly care" and in standardizing data collection are now taken for granted by the facilities and new members of MaiKhanda team[27]. The MaiKhanda's District Coordinators are posed to play a central role in working with the Ministry of Health officials to help get the key changes developed and tested in our facility collaboratives made a permanent part of the health care system. This will include changes in policy, guidelines, forms, job descriptions, and ongoing training of health care workers. They will have to play a bigger role in our QI efforts as we move from testing to implementation[27].			Most changes are not being reliably implemented and have not been fully integrated into local, district, or national structures of the Malawian health system. This was partly due to the randomization design , but will now require a deliberate design at all three levels[30]Some form of scale-up across the districts may be the only way to ensure that the project meets its stated aim of decreasing maternal and neonatal mortality in the districts. The current patchwork of interventions may not deliver the result that could be achieved from a more integrated approach[30].	

	Intervention Periods	Baseline period (June 2006-January, 2007)		Pre-intervention period (February, 2007-September 2008)			
	Years	2006		2007			
	Quarters	Q3	Q4	Q1	Q2	Q3	Q4
B	**Outer context**						
1	External policies and incentives	Road Map proposed by the MoH aims to develop and implement appropriate behaviour change intervention and advocacy strategies to address maternal and newborn health issues, to conduct maternal death reviews and clinical audit and strengthen health facilities at all levels to ensure adequate coverage with functional CEOC and BEOC facilities [12].	At the time of the baseline survey in 2006, national policy was that at ANC women were requested to bring with them during delivery, a clean plastic sheet to lie on, a clean razor blade to cut the cord, and two chitenjes, one to cover herself and the other to wrap the baby.Women who deliver in HFs adhered to this request[2]. This is however seems to have phased out by the start of the intervention period (although no official documentation is available on this)			The political environment in Malawi has been quite stable over the past year. With the government bid to end corruption there is confidence in money being used optimally thereby providing greater confidence among donor agencies to invest in the health sector in Malawi. A good example is that of SWAp along with the EHRP (Emergency Human Resource Package) being implemented across the health sector with commitment from different stakeholders [12]	MaiKhanda program implementation was affected by competing programs of their partners. For example , DHOs in both Kasungu and Lilongwe are also working with other partners like MSH and JHPIEGO that are involved in Quality improvement and other health related MNH programs. It was difficult for the partners to concentrate fully on MaiKhanda program activities [14]
2	Health systems readiness	Malawi has almost double the recommended minimum number of Comprehensive EmOC facilities (1.8 facilities per 500,000 population) while only 2% of the recommended number of Basic EmOC facilities (0.1 facilities per 500,000 population) There is no equity with respect to geographical distribution. The met need for EmOC is about 18.5 %, which is far below the UN recommended level of 100%. Only 2.8% of all expected deliveries are by caesarean section which is below the recommended minimum of 5%. Quality of EmOC services is generally poor as seen by a case fatality rate of 3.4% which is much higher than the UN recommended level of less than 1% [13]	928 people were working in the 29 Health Centres surveyed at baseline. Of these, only 126 (14%) people are clinical staff and rest (86%) non clinical staff. Of the clinical staff, 37 (29%) are COs/MAs and 89 (71%) are Nursing Officers; and of the 802 non-clinical Health Centre staff, 509 (63%) are HSAs, 143 (18%) are Patient Attendants and 150 (19%) are other staff [4].			Maternal death reviews were conducted regularly in all 3 districts (Kasungu, Lilongwe & Salima). tandard Safe Motherhood (SM) treatment protocols for the major emergency obstetric complications were available in nearly all maternities in the three districts[2]. While, there is an expectation from the MoH that Health Facilities will complete all their own MDRs, in reality the facilities are presented with many challenges. For example, Kamuzu Central Hospital (KCH) experiences a large number of maternal deaths, and often there are not enough personnel available to complete each MDR [12].	

	Intervention Periods	Pre-intervention period (February, 2007-September 2008)		
	Years	2008		
	Quarters	Q1	Q2	Q3
B	**Outer context**			
1	External policies and incentives		The global rise in fuel prices has also had a negative impact on transport for project activities. Fuel prices have gone up thereby impacting on the current budget [17]	Other MNH projects were being implemented in the health facilities included: Antenatal Care Improvement, Mother to Mother/ CTC, IMCI, CBDA supported by MSH, Advanced Life Support Obstetric Course (ALSO), CESTA which Provides refresher training to midwivesand allowances for locum shift midwives, home-based care by MSH, Reproductive Health sponsored by UNFPA, Infection prevention, PMTCT Programme [7]
2	Health systems readiness	There is a lack of capacity within CHAM or Govt hospitals for facilitation [15]	There were some changes in leadership on the part of the Ministry of Health. This did not have a negative impact on the project. The new minister is very supportive of MaiKhanda project including the Principal secretary [17]	LLW has a population that is nearly 6 times the size of a normal district. The HCs have on average 1 0,000 additional population to deal with thereby increasing workload[6].

	Intervention Periods	Early intervention period (October 2008-September 2009)			
	Years	2008	2009		
	Quarters	Q4	Q1	Q2	Q3
B	Outer context				
1	External policies and incentives				
2	Health systems readiness	While facility QI teams and other health workers are keen on improving the quality of care, in many facilities (e.g. in Mlare and Bwaila) overcrowding in nursery and delivery areas and shortage of equipment and supplies (e.g. lab reagents) are major constraints [19]			

	Intervention Periods	Intensified intervention period (October 2009-December 2010)				
	Years	2009	2010			
	Quarters	Q4	Q1	Q2	Q3	Q4
B	**Outer context**					
1	External policies and incentives		Health facility deliveries have risen rapidly across the whole programme area from a low of 36% in the middle of the 2007/8 rainy season to a high of 64% in August 2009 for all three districts combined. This increase might be due to government policy to ban all TBA deliveries, improvements in the care offered at facilities, increased demand stimulated by MaiKhanda Women's Groups and possibly the effect of our village volunteers (community surveillance) encouraging women to deliver at a facility[31].		At national level, a number of developments have combined to slow down the advocacy and stakeholder engagement efforts. , The appointment of a new Minister and Deputy Minister of Health and the transfer of the Principal Secretary of Health meant that we have to start rebuilding relationships at senior levels of the Ministry[9]	April to June consistently has the lowest mortality. It's speculated that this is a combination of a natural reduction in number of births and the ease of getting to a facility in good time due to relative ease in passing the roads in dry season[32].
2	Health systems readiness		Health centres in Salima appear on average to be more women-friendly than other districts. This resonates with other studies (HF reources survey) which indicate the existence of a relatively better health system/service in Salima[8]. NND vs MD: The focus on reducing NND and NNCFR is currently not as strong as MD because facilities have chosen to focus on MD. This is because HC don't have the equipment that helps them with NND – there are no ambu-bags, etc., but even with bags,there is still need to create a system in place before hand [31]		Huge turnover of staff, so there isn't a consistent cohort e.g. a new QI team has only just been established at Bwaila[9].	

	Intervention Periods	Baseline period (June 2006-January, 2007)		Pre-intervention period (February, 2007-September 2008)			
	Years	2006		2007			
	Quarters	Q3	Q4	Q1	Q2	Q3	Q4
B	**Outer context**						
3	Resource availability						
a	*Human Resources*					Facilities themselves perceive there to be difficulties in motivating team members and their colleagues to participate in QI work. Every one of the nine CEmOC facilities noted that a lack of material and human resources caused problems with QI implementation[12].	
b	*Material and resources*	Screens for delivery were available in 81% of the health facilities surveyed at baseline. Companionship during labour was provided in only 30% of the surveyed facilities. Bedsheets to cover women during labour were available in 35% of the facilities. Only 40% of the expected deliveries were being attended by a skilled birth attendant [2]. Only 31% of the facilites in the baseline survey had all the materials functional[2]				Inadequate materials and supplies for obstetric care continue to affect the operations of many health centres and facilities. Hospitals and health centres continue to experience some problems in adequate materials and supplies. One of the contributing factors is inability of DHMT to include critical needs in their District Implementation Plans[14].	
4	Organization culture/climate					All the facilities, with the exception of Mlale and Salima reported encountering problems with their QI implementation. These problems ranged from poor infrastructure to lack of personnel. A common response was resistance to change by other staff members, and a lack of morale and commitment of staff[12].	

	Intervention Periods	Pre-intervention period (February, 2007-September 2008)		
	Years	2008		
	Quarters	Q1	Q2	Q3
B	**Outer context**			
3	Resource availability			
a	*Human Resources*	High turn over of staff continues to affect the quality and progress of QI activities in most facilities that MaiKhanda is working. Majority of the health workers that were trained in QI have either left or been transferred thereby creating a gap[17].	At district level, there was a high turn over of facility staff. This affected the implementation of most QI activities. Two facilities were hardest hit by staff turn over. For example Salima hospital and Likuni faced a lot of challenges. A number of staff that were trained in QI methods resigned and some went for further studies thereby creating a knowledge gap [17]	HF resources survey of 29 HCs showed only 14% of the staff are clinical with a majority of them nurses (71%).Availability of nursing staff at night is low in Salima (31%) [4]. Shortage of staff led to an increased workload which did not give the QI team members enough time to meet regularly and discuss QI activities. Another similar reason for the shortage of staff for QI activities was the high turnover among the members which hindered progress in QI work.
b	*Material and resources*	Lack of adequate blood supplies also pose a challenge for effective emergency obstetric care. Kasungu and Salima district hospitals continued to face some problems related to blood supplies. Lack of drugs and supplies continue to pose a challenge on the quality of care being provided to women[17].		14 (48%) health centres have all the materials available but only 9 (31%) health centres have all the materials functional. (52%) of the health centres have both electricity and water supply, 4 (14%) have electricity only, 3 (10%) have water supply only while 9 (31%) of the health centres have nothing.These are mainly the rural health centres which also have poor roads. However, most of the health centres in Salima are also rural, but more than half of them have road access and all of them have access to water supply and electricity[4].
4	Organization culture/climate	Of the facilities, Likuni Mission hospital and St.Gabriel's seemed to be the most active in implementing the change packages using the improvement cycle. Both are CHAM facilities [7]	There are differences between the districts on a variety of determinants. LLW is worse of on issues related to happiness, job interest and conflict with patients (all of them are statistiaclly significant when compared with other districts).	An environment fostering innovation and autonomy by staff to carry out improvement work may be lacking at times, meaning that it may be more difficult for them to test out new ideas using the PDSA cycle. In general, it appears that the morale of staff is good with the majority of respondents reporting that they are happy in their work, that they get along with their colleagues and can turn to them for work advice / support, and that they don't often have conflicts with other staff and patients.while they do report being appreciated by their supervisors, they are often not consulted on work related issues. However, urban staff are worse off than their counterparts in rural HC with regards to will and motivation, experiencing conflict with other staff, accepting additional job responsibilities, having freedom to take decisions about their work, encouraging subordinates at work and being able to talk freely to their immediate supervisor[6].

	Intervention Periods	Early intervention period (October 2008-September 2009)			
	Years	2008	2009		
	Quarters	Q4	Q1	Q2	Q3
B	Outer context				
3	Resource availability				
a	*Human Resources*	Critical shortage of neonatal clinicians and neonatal care skills in facilities is reducing attention to newborn care, e.g. Kasungu has not had a pediatrician for 10 years[19]			Constant shifting and re-allocation of staff from maternity to other wards slows down the pace of QI activities, Posting of staff to other facilities has also continues to disrupt QI initiatives[25].
b	*Material and resources*				At health centres supplies of oxytocics and antibiotics have not improved since 2007: half are without these essential drugs at any one time. Most of the 62 health centres provide only 1 or 2 of the 6 signal functions of basic obstetric care and some not providing any at all [31].
4	Organization culture/climate		Most health centres not following up effectively on QI projects they have committed to undertake. Frequent transfers of health facility staff (QI team members, data personnel, etc) leading to a weak organizational memory about the interventions. Frequent absences of health facility staff for various reasons, but especially for workshops. Participants in data learning sessions keep changing, making it difficult to maintain progress from session to session [19]		

	Intervention Periods	Intensified intervention period (October 2009-December 2010)				
	Years	2009	2010			
	Quarters	Q4	Q1	Q2	Q3	Q4
B	**Outer context**					
3	Resource availability					
a	*Human Resources*		Staff turnover is a huge problem – it takes about 9 months for local sustenance of the intervention. Although there's only an 11% rate of staff turn over, it's usually the high functioning people that get moved [31]		There is a growing demand for skilled delivery against limited supply of quality services. With more women delivering at health facilities, the capacity of health facilities is severely tested [9]	
b	*Material and resources*					
4	Organization culture/climate		Only 50% of the staff said that they would recommend a friend/relative to deliver at their facility[5]			

	Intervention Periods	Baseline period (June 2006-January, 2007)		Pre-intervention period (February, 2007-September 2008)			
	Years	2006		2007			
	Quarters	Q3	Q4	Q1	Q2	Q3	Q4
C	**Inner Context**						
1	Governance and oversight	There was discussion in particular of the need for the workplan to reflect the attempt to create a single, shared model of QI and all activities to be owned and genuinely led by the team on the ground, not by overseas organisations. It was agreed that a far greater degree of clarity is needed about visits to Malawi by overseas individuals, particularly TA roles. All visits – their timing,purpose and agenda need to be driven by local needs, rather than by what suits the diaries or interests of TA providers[10].		Malawi team to manage at operational level and the experience the team brings to programme development and identifying budget implications, to be recognized. A better balance to be sought for programme planning and budgeting between operational and strategic level and good communication between the two levels needed[11].		It has been observed that there are misunderstandings concerning the integrated model, and with regards to which partner should take the lead in QI. This has resulted in the perceptions that there is no ownership of the model. This has led to confusion in the facilities, reflected in the confusion regarding MaiKhanda identity. Identiity: There are discrepancies in how the partners of MaiKhanda identify themselves, for example some staff members introduce themselves as MaiKhanda, whereas others affiliate themselves with their parent organization. This inconsistency in the identification of MaiKhanda as one team results in confusion at the facility level[12].	A full-time Programme Director (PD) within Malawi, who will hold a single programme budget and will deploy it in line with a single integrated workplan was agreed . IHI was the principle technical advisor for a single coherent model and approach for the programme as a whole which integrates the range of methods offered by the partners within the consortium[13]. LSTM withdrew from the Consortium and all MNH activities will be taken up by IHI and other partners [14].
2	Donor role/partner relationship		The language of partnership is lacking between the different partner organizations and therefore, while MaiKhanda has a very strong and committed staff working on the ground, the commitment of these staff members to their parent organizations affects their functioning as a coherent team. The implications of such a disintegrated approach are more subtle. It is not reflected in the day to day activities performed interchangeably between the staff members rather, the implications are more intrinsic. The various team members during their interaction with the evaluation team, have expressed increasing amount of frustration in performing their duties as a 'team'[12]. Clarity regarding the on-going involvement of TAs is needed and should also be aimed at developing the capacity of the Malawi team to lead and develop the programme in Malawi[11]. MaiKhanda team should be asked to work on this (integrated model) first and the importance of their local experience must be given due weight relative to the theoretical positions of the consortium organisations[12] The donor role as 'partner' as opposed to donor adds to the complexity of the consortium. There is a lack of clarity among the consortium members as to what the donor's role as partner is or might be in the future. In addition their perceived close affiliation to one of the consortium partners is likely to have had an indirect influence on levels of confusion or frustration among staff on the ground[12] .		The team on the ground is not being given adequate breathing space to develop their full potential given the differences of opinion among the senior members of the consortium on the way the programme needs to be implemented. This has caused a lot of grief and frustration among the forbearers of the programme on the ground affecting staff morality and level of functioning[12]. Discussions were held with partners to see if there would be better ways of implementing the integrated model. However this did not work and resulted in the exit of LSTM and LATH[14]	Visits by overseas clinicians were poorly coordinated, despite best efforts in Malawi; and were not delivering the best benefit. It was becoming clear that the lack of integration was leading to waste and duplication of effort on the ground, that it was undermining the Programme Manager role and was demoralising for staff. All partners to re-affirmed that they would not be working directly through their respective programme officers i.e shift from "decider" to "advisor" for partners [13].
3	Network and communication			On-going communication between all partners still needs to be improved particularly in relation to the development of workshops and between facilitators and implementation team when the workshops are in progress, especially when changes are made to the programme.More prompt communication was requested from LATH regarding response from partners on financial matters[11].	It has been observed that partners are deviating from what was written in the proposal. In addition, as one MaiKhanda staff member explains "There have not been any consultations or even communications with the other partners" about these deviations. There have been a number issues being raised with regard to communication within MaiKhanda, for example appropriate staff members are not receiving the correct information[12].		Communciation remained a major problem within the consortium. The partners continued to communicate directly with their responsible Program Officers without involving the Program Manager. However this problem has been addressed by the proposal to recruit the Program Director. It is envisaged that once the Program Director is in place, all program officers will be communicating through the Program Director[14]
4	Capacity within MaiKhanda					MaiKhanda does not have officers responsible for MNH and QI activities at district level. With the rolling out of QI/MNH activities in to health centres, it was extremely difficult for the available officers to conduct mentoring and coaching visits to all facilities and follow up on the teams projects. There is need to consider having district based QI/MNH officers who could provide ongoing support to QI teams at district level. Training activities for MK staff: Life saving skills for MNH officers, Imporvement advisory training for QI program, Training in financial & accounting system. Super imporvers trained in Jul'07 & Improvement Advisory Board established [14].	

C	**Inner Context**			
1	Governance and oversight	A unified workplan and budget managed locally by the Program Director agreed. This was to strengthen local ownership and proper budget tracking and monitoring.Changes in programme management and the pulling out of LATH created a gap in logistics and administrators role that was quite challenging[17].		There are regular Monday morning meetings in each of the 4 offices for all the staff in the office. In addition, there are monthly (on average) meetings involving program and admin staff of the Central Offices plus the DCs representing the District Offices. However, there are no regular meetings of what might be called the management team.[19].
2	Donor role/partner relationship	THF conforms to adopting the IHI approach as the overarching framework[13]. IHI endorses a single design for improvement activities will be more appropriate and more effective in pursuing the program's primary aim, since the prevoius integrated model failed to work as effectively as needed in a coordinated manner, threatening sustainable progress and increasing the burden on local program staff and participants[13]The 'silent partner' Cincinnati Children's Hospital and Medical Center (CCHMC), was to be recognised as a true partner in the project in the revised programme strategy. . Additionally, a new name for the programme is needed[15].		
3	Network and communication	WCF expressed concern that they were not fully informed on what was going on in the Programme – i.e. they have not been involved in the facility and community linkage. Secondly, they had not been informed about what activities the Technical Lead had been doing[15].		Partners we work with in the districts do not always see us as one project because MaiKhanda staff from the districts are not involved in QI activities. Our friends from the central office just come into the district and we sometimes just hear of events happening via the partners.MaiKhanda central office needs to fully engage district staff in QI so that they know what is happening[19].
4	Capacity within MaiKhanda		MaiKhanda puts great emphasis on staff development and in the budget, there is K1, 000,000 set aside for staff development. In addition, senior staff are on the IA (Improvement Advisor) course supported by IHI. A draft training plan with areas of need was developed. Recruitment of 3 additional District Quality Improvement Officers proposed and approved[18].	

	Intervention Periods	Early intervention period (October 2008-September 2009)			
	Years	2008	2009		
	Quarters	Q4	Q1	Q2	Q3
C	**Inner Context**				
1	Governance and oversight	The absence of quality improvement staff in the districts and the generally weak communications between district staff and the central office quality improvement team,has been a constraint ot MaiKhanda district-focused approach. In the eyes of many of district level stakeholders it has looked as though there are two separate organizations in operation -- a community-based one and a facility-based one[19]	LATH Trust/The Health Foundation Consortium officially becomes MaiKhanda [20]The MaiKhanda Senior Management Team (SMT) was formed.Creation of District Management Team [21]. District Coordinators to take up role as District Managers and be responsible for all activities within their district [19]. MaiKhanda held a highly successful Strategic Planning Workshop in January to develop a 3-year Strategic Plan (2009-2012) and an Annual Workplan and Budget[20].		Decentralization – there is growing unity of among functional teams across districts and growing unity across units within districts. The increased capacity of the FI team with district presence and growing QI knowledge and skills. Growing operational stability is increasingly allowing the team more time and space to focus on advocacy and stakeholder engagement[25].
2	Donor role/partner relationship	Structure within MaiKhanda has been confused by power relationships.There have been occasions when partner representatives have arrived in Malawi and have acted in some ways as line managers of Malawian staff, saying things such as "You must do this activity"[19].			The design of the RCT is a constraint on the spread of the project[30].
3	Network and communication	This has been a challenge between us districts and with central office. Many problems have arisen due to lack of communication, for example, on allowances – each district has chosen its own levels of allowances and this is confusing our partners who think we are different organizations. Communication from central office has also been a pr*oblem – "many times you just get a call or an email saying things have now changed and you are to do this or do that without any written formal communication stating the policy; written communication is important"*[19].			Scheduling and maintaining appointments for follow up at HCs is often difficult- this is largely due to limited staff on-site. • Transport: transport continued to be a challenge due to the age of most of the motorbikes and the limited number (and poor condition of one) of the four-wheel drive vehicles[25].
4	Capacity within MaiKhanda			"Communicating what we are learning": a Skills Building Workshop for MaiKhanda Staff [20]. Significant focus on professional development for both existing team members and the new recruits. Orientation of new FI members and ongoing on-the-job coaching and mentoring of the entire team by IHI faculty[23].	QI training is done in FI Team Monday meetings[24]. Recruitment of district-based FI officers has stabilized and improved focus on QI activities with more frequent coaching visits to QI teams, which is critical in influencing QI application on the ground [25]

	Intervention Periods	Intensified intervention period (October 2009-December 2010)				
	Years	2009	2010			
	Quarters	Q4	Q1	Q2	Q3	Q4
C	**Inner Context**					
1	Governance and oversight					Concern is raised over who ultimately makes decisions. Agnes suggests that the Salima Proposal and the Communications Strategy die because the decisions are ultimately made "over there" [32].
2	Donor role/partner relationship					
3	Network and communication				Communications workshop held in June 2010 together with a learning event. This generated a lot of energy and enthusiasm among the different teams (district and central). The district teams worked together to prepare their district quarterly reports[W43].	
4	Capacity within MaiKhanda		The MaiKhanda program staffed to provide reasonable QI support to the facility improvement teams. CEmOC's multiple visits planned, additional accelerated support to Bwailia with multiple visits each week to their QI teams. MK staff completion of the Improvement Advisor Development program prepares them to assist Senior Management in QI. The data capabilities of the FI team are strong. Efforts to provide the data analysis to guide FI officers[27].	Custom-designed week-long management and leadership development workshop attended by 16 members of the district and senior management teams in June [9].	There have been changes in the FI team members that have effected our continuity in supporting front-line QI teams [9].	

	Intervention Periods	Baseline period (June 2006-January, 2007)		Pre-intervention period (February, 2007-September 2008)			
	Years	2006		2007			
	Quarters	Q3	Q4	Q1	Q2	Q3	Q4
D	**Individual characteristics**						
1	Knowledge and beliefs about the intervention			Facilities themselves perceive there to be difficulties in motivating team members and their colleagues to participate in QI work. Every one of the nine CEmOC facilities noted that a lack of material and human resources caused problems with QI implementation[12].		Some facilities felt that they were being pushed too much by QI activities, and felt frustrate at being steered too much into different directions. One facility reported that they were not able to keep up to date with everything, and spread the changes to all the staff. This was because the changes are happening too quickly, in addition to the fact that there were too many changes occurring as part of the QI programme [12].	With the introduction of women friendly standards, delivery of pregnant women at facility level has improved. However the hospitals are also seeing an increase in the number of maternal deaths. Perhaps this is so because hospitals are now able to attend to cases that would have delivered at TBAs or Home[14]
2	Self efficacy					In the facilities there is not the same level of awareness of neonatal deaths as there is for maternal deaths, and this may result in less confidence in previous neonatal data from the facilities[12]	Quality improvement activities especially criterion based and maternal deaths audits have proved very effective in the provision of quality obstetric care. With the introduction of women friendly standards, delivery of pregnant women at facility level has improved[14].
3	Motivation		Collaborative workshops allows for regular interaction between the facilities, and a number of facilities use the QI work as an opportunity to meet with clinical staff from other facilities. The motivations for individuals to become involved in QI range from being in a team, to learning from others to being able to identify problems and find the solution. [12].			Minority of facilities felt that they were being pushed too much by QI activities, and felt frustrated at being steered too much into different directions. One facility reported that they were not able to keep up to date with everything, and spread the changes to all the staff because the changes were happening too quickly[12]	
4	Team work and stability	In general, a team was to be comprised of representatives from the various provider roles within the facilities, for example, QI team members typically included a Matron, Nurse-Midwife, Clinical Officer, Lab Technician, and Data Clerk. However, a few of the QI teams, such as the team in Kasungu, developed from the existing MDR committee. It was observed that QI team members have a high level of commitment to the QI work, for example often the team members will attend a QI meeting even if they have been on duty the previous night. Feedback on the QI programme has been mainly positive, with a number of team members reporting personally that they really enjoy the QI work[12].	A number of teams are composed of a wide range of facility staff representatives including: the Medical Officer, the Chief Clinical Officer, the Matron and the driver on duty. This would ensure that all the perspectives of care within the facility are considered, including transportation issues[12].			QI programme has facilitated the formation of a network of facilities that are able to share ideas, knowledge and also resources. One such example is that of a situation that arose when Mtundu were out of antibiotics, and Nkhoma shared with them some of their near expiration antibiotics[12]	
5	Leadership						

D	Individual characteristics			
1	Knowledge and beliefs about the intervention		About 52% of the respondent felt that the QI systems fits into their current system of work but with some effort. This was mainly because of the shortage of staff. A majority of the respondents felt that QI qas a little bit demanding , but worth the effort.QI teams felt that they could improve the quality of care at their health facility, despite the many constraints of material and human resources. This was also one of the reasons motivating them [7].	
2	Self efficacy	As part of clinical skills building for QI teams, MaiKhanda identified and engaged clinical advisors to support the QI teams on various issues related to improving quality of care. These sessions revealed the need for ongoing support for neonatal resuscitation as many health workers lacked adequate knowledge about basic resuscitation measures[17].		All respondents from CEmOC survey valued visits by MaiKhanda staff. 52% of them suggested that MaiKhanda should visit their facilities twice a month; 44% suggests that once a month visit would be adequate for their facilities. Shortage of staff led to an increased workload which did not give the QI team members enough time to meet regularly and discuss QI activities. “Since we are short staffed sometimes it becomes difficult to follow and do a protocol
3	Motivation		Improvement projects are instilling some kind of discipline among QI teams such that they take it as their responsibility to ensure that women coming to them for care do not die of complications that could be avoided[17].	An important motivating factor was the new knowledge the participants gained from attending QI workshops as well as the learning they derived from the QI work of other hospitals[7].
4	Team work and stability			One of the primary reasons motivating the QI teams was that they felt that they could improve the quality of care at their health facility, despite the many constraints of material and human resources. Stability of QI team members does seem to have an effect on QI work being carried out in the health facilities. Challenge of a lack of teamwork is present and is linked to the human resource crisis. Because human resource is reallocated whenever necessary, so each time new staff come in they have to be oriented.[7]
5	Leadership		70% of the quality improvement team stopped carrying out their projects due to lack of support from senior management and lack of team spirit [17].	On average, staff enjoy a ‘good’ relationship with their superiors. Leadership does not seem to be inclusive as staff are less likely to be asked by their supervisors for suggestions and comments on work related issues. Staff also complain of a lack of supervision by their supervisors[6]

	Intervention Periods	Early intervention period (October 2008-September 2009)			
	Years	2008	2009		
	Quarters	Q4	Q1	Q2	Q3
D	Individual characteristics				
1	Knowledge and beliefs about the intervention				
2	Self efficacy			Data clerks in the data collaborative are building on the run charts initiated in the Data Workshop in June 2009 and they are showing a wider understanding in interpretation of run charts[23]	QI Skill: health centres are showing great improvements in their day to day work, especially in the area of data capture and use; for example, during the last series of HC learning sessions, 80% of health centres captured and used family planning data alongside data on deliveries – demonstrating a growing grasp of the whole systems approach. Good progress noted with Salima team- teams able to present progress using run charts [25]
3	Motivation	Younger staff (18-30 year old) are consistently better off than others when it comes ot motivation. However, older staf (i.e above 50 years) seem to be more content and settled in their work lives and more comfortable with authority. Staff in CHAM facilities are better off in terms of motivation, and coming up with new ideas at work. Nurses are worse off than medical assistants in maintaining relations with other staff members. MAs are worse off when it comes to knowing the patients by their name. The degree of engagement of nursing staff with the care being provided is much more. Motivating factors: helping other people, team work, personal gains such as electricity in rural area.Demotivating factors: lack of support systems(drugs, staff), high workload, lack of career progression. On average, staff are interested in their work and often willing to work beyond their job responsibilities if required, but are sometimes lacking in motivation, partly due to the fact that they are almost universally completely dissatisfied with the salaries they receive [6]. Staff motivation survey: Lots of anecdotal evidence of poor resource availability in the facilities and how it affects staff motivation and morale[6].			Concerted support to QI teams in facilities yields good results. Frequent coaching visits to QI team in the facilities motivates them. There is now an increased dependency on data for decision-making amongst QI teams in the facilities, Health Centre and CEmOC Collaboratives have changed the mindset of many teams on the provision of quality care for the better as they continuously learn from each other and develop new change ideas for testing [25].
4	Team work and stability	130 members have once been members of QI teams from mid-2006 to mid-2008 (till the CEmOC survey was conducted), with 88% still retaining their association/involvement in the QI team. Another 11% are no longer QI members. Although only 15 of the QI team members dropped out during the survey period, there is anecdotal evidence to say that they were the most dynamic within the team and exhibited some exemplar leadership qualities[7]	Staff do have a network within their facility for support and discussion of problems while network between facilities is not that well established. The overall picture is that staff do get on well with each other, can discuss problems with colleagues, encourage each other and have good relations with their superiors. However, there needs to be more links between facilities.These informal links are necessary especially for the spread of QI intervenions[6].		
5	Leadership				2-day learning session for leaders from Lilongwe health areas to orient them on QI and strengthen relationships between them and different levels of management at DHO level[25].

	Intervention Periods	Intensified intervention period (October 2009-December 2010)				
	Years	2009	2010			
	Quarters	Q4	Q1	Q2	Q3	Q4
D	**Individual characteristics**					
1	Knowledge and beliefs about the intervention					
2	Self efficacy	Only 42% of the HF staff felt that the care they provided was ideal. 37% of the repondents (mainly nurses) felt that women should labour in the position of their choice. Knowledge in CEmOC facility staff was greater than BEMOC facility staff[5]. Provider knowledge survey	The teams have tremendously improved in the way of reporting which showed a maturity and understanding in what they are doing. The format of presentations, the language and the graphical presentation all displayed an advanced level of understanding of the quality improvement in all the health centers[28].			
3	Motivation					
4	Team work and stability					
5	Leadership				Clinical and senior leadership was lacking in most health facilities. In places where leadership was strong and supportive of MaiKhanda's interventions, visible improvements were made- for example Salima and Nkhoma hospital[9].	Salima district, mainly due to its activated leadership, is poised for rapid spread and could be a catalyst for the other district managers to follow suit[30].

	Intervention Periods	Baseline period (June 2006-January, 2007)		Pre-intervention period (February, 2007-September 2008)			
	Years	2006		2007			
	Quarters	Q3	Q4	Q1	Q2	Q3	Q4
E	**Implementation process**						
1	Organization of work.			A better balance sought for programme planning and budgeting between operational (local team) and strategic level(TAs) and good communication between the two levels on programme management issues[11]. The size and cost of QI workshops has far exceeded than expected or planned.The on-going need for large, centralised workshops needs to be reviewed and guidance from the Malawi team sought in this respect.[11].		PM and the local team to become much more active "commissioners" of TA. All visits should be added to workplan so that someone's presence can be properly managed and utilised by the local team[10]. Focus on district and central hospital(s) (while continuing to work with smaller hospitals at same time)[13].	District based QI/MNH officers to provide ongoing support to QI teams at district level[14].
2	Engagement(stakeholder)	The programme was initiated in Malawi in February 2006 with a stakeholder meeting between MaiKhanda members and key stakeholders from the MoH and health facilities to develop the programme plan of action[2].		The RHU (MOH) was keen to have other districts involved in scaling up MaiKhanda programme.However, it was explained that with the current level of human resource in the programme in Malawi it was difficult for MaiKhanda to take on more work in other districts[11].		The institutional environment has been quite conducive to the implementation of QI activities. Members of the RHU have been involved during our programme management board meetings providing valuable inputs. The February workshop was attended by the Honourable Minister for Health. She appreciated the efforts of the programme and encouraged the participants to be actively involved in the QI work being conducted by MaiKhanda[12]	Involvement of the district stakeholders in the Collaboratives remained poor. There was little involvement of district based stakeholders in QI activities as such some problems requiring intervention by DHMT or District assemblies[14].

E	Implementation process			
1	Organization of work.	Difficult to develop 'one plan one budget' as the initial arrangement has some elements (i.e. travel, partner expenses) that are still controlled by partners outside the MaiKhanda. Current Budget lines still reflect partner's expenditure. There is no clear linkage between in-country plans and partners activities[18].		Expanded Facilities Intervention Team – 3 new members have been recruited to the team (one each for Salima, Lilongwe and Kasungu and one to manage Bwaila[23]. A new accounting package, Sage Pastel Evolution, was procured, installed, and is operational, replacing the LATH accounting package, PAS and a new chart of accounts was developed. The Board of Trustees and the Partners Board both had successful meetings.[22].
2	Engagement(stakeholder)	Non-active involvement of stakeholders; need to share data with other stakeholders[15].	Four meetings were conducted between MaiKhanda and members of the district health management and hospital management team in Kasungu, Lilongwe and Salima.District Executive Committee (DEC) meeting in Salima. 30 ADC (Area Development Committee) meetings to sensitize stakeholders re the program [17]. MaiKhanda also engaged with the Safe Motherhood Taskforce, Quality Assurance Task Force[18].	

	Intervention Periods	Early intervention period (October 2008-September 2009)			
	Years	2008	2009		
	Quarters	Q4	Q1	Q2	Q3
E	Implementation process				
1	Organization of work.			Ongoing management improvements focused on strengthening the district as a unit of implementation[23].	All districts reported a constant flow of funds to the district; this has brought satisfaction to the team as they have managed to meet their program targets[25]. MaiKhanda vision of the future to use districts as the unit of planning and implementation. MaiKhanda to have an integrated design and management in each district. District Coordinators presented a consolidated report (for FI, CI & M&E) for their districts[24]
2	Engagement(stakeholder)	Weak statutory coordinating mechanisms (e.g. RHU/MoH, local government institutions such as District Executive Committees and community level development committees) force MaiKhanda and other NGO partners to spend too much time coordinating (though not always successfully as without mandate)[19].	Participation of a significant number of key stakeholders in the strategic planning meeting including RHU (MoH); Save the Children; World Health organization; MaiMwana; Mother to Mother; Kasungu, Lilongwe and Salima DHOs; and Nurses and Midwives Association of Malawi. Ongoing dialogue with Sakaramenta, NICE (National Initiative for Civic Education), White Ribbon Alliance MNH Advocacy Campaign for Aspiring Members of Parliament, Radio Zodiak[20].	MaiKhanda involved in the development of District Implementation Plans (DIPs) and budgets for the year July 2009 to June 2010 for Lilongwe, Kasungu and Salima districts for the first time[22]. Frontline staff continue to be challenged by the huge expectations fueled by the extreme poverty in the areas MaiKhanda work and the “handout syndrome” that has become so prevalent in Malawian society over the last 10-15 years[23].	Director presented a detailed written briefing to the new Minister of Health followed by an introductory meeting. Minister convinced that MaiKhanda’s unique approach to maternal and newborn health has a lot to offer to the country. He challenged MaiKhanda on how they intended to take this small-scale “pilot” and apply it to the rest of the country[25].

	Intervention Periods	Intensified intervention period (October 2009-December 2010)				
	Years	2009	2010			
	Quarters	Q4	Q1	Q2	Q3	Q4
E	**Implementation process**					
1	Organization of work.	Scheduling and maintaining appointments/meetings for follow up often difficult, largely due to limited staff on-site. Together with Partners, MaiKhanda has also devised some solutions to the major transport problems, which are about to be implemented[25].				
2	Engagement(stakeholder)		Met with DHO(Kasungu) and one-on-one meetings with QI team members to find out reasons for absenteeism from QI meetings. Explained to them the importance of such meetings. Immediate plan is to meet with the new DHMT for support of the work[28].		At national level, a number of developments have combined to slow down the advocacy and stakeholder engagement efforts. The appointment of a new national Goodwill Ambassador for Safe Motherhood, which took a few months, meant we had to put on hold some of the initiatives that had been in the planning pipeline, e.g., discussions on organizing a major national event to share lessons learned/best practice. Equally, the recent appointment of a new Minister and Deputy Minister of Health and the transfer of the Principal Secretary of Health (new one yet to be appointed) means that we have to start rebuilding relationships at senior levels of the Ministry [9]	

	Intervention Periods	Baseline period (June 2006-January, 2007)		Pre-intervention period (February, 2007-September 2008)			
	Years	2006		2007			
	Quarters	Q3	Q4	Q1	Q2	Q3	Q4
E	**Implementation process**						
3	Execution						
a	*Fidelity*		Size of the intervention: The intervention was designed taking into consideration the normal district structure. However, the MoH had LLW thrust upon us as they considered it as one of the districts which needed more attention. LLW is as big a Salima and Kasungu put together. While, these were events beyond MaiKhanda's control, it has had significant management and budgetary implications. Although there was an excellent team in LLW, the required scope of work was stretching them to their limits[12]	A number of innovative ways have been developed by the facilities to generate ideas for implementation. These developments may be considered to be a positive step towards sustainability, in that facilities are developing their own way of doing QI in a way that works best for them. While the intervention(s) were designed within the framework of a RCT, the 9 referral CEOC facilities was not meant to be part of the randomization (control) process. It was decided to implement the intervention across the 9 facilities as a health system strengthening exercise[12].		Turnover of staff makes organizational learning, memory weak. Significant confusion, parallel work and very heavy reliance on numerous training sessions (low reliability)[13]	
b	*Intensity*					A neonatal expert meeting was conducted by content experts in neonatal and paediatric care along with representatives from the facility QI teams[12].	
c	*Dosage/Coverage*	Interventions started in 2006 with a focus on the 9 CEmOC facilities in the 3 districts. An initial target of 50% reduction in mortality in the first year of QI work was set[3]					The key activities implemented were: the rolling out of the 90 day project on quality improvement work to health centers; establishment of an improvement advisory team; super improvers training; implementation of a Data collaborative session; Follow up on CBA, 4 MDRs, facilitating implementation of women friendly standards for health centers; training on life saving skills for all the CEMOC facilities and health centers; learning session on care of the neonate[14].

E	Implementation process			
3	Execution			
a	*Fidelity*			The question was posed if MaiKhanda should start providing essential materials used during hospital training (in addition to the $30,000 worth of supplies provided quarterly by Cincinnati Children's). This cost could be budgeted into Phase Two. The implications of being a provider need to be further reviewed as the current policy is to provide the evidence to allow the hospitals to ask the Ministry, or budget it in themselves. This was not approved by the PMB[19].
b	*Intensity*			In total there were 162 ideas throughout the 10 CEmOC and that most (87) of these ideas were related to clinically excellent care. Only just over half (56%) of the total number ideas were tested with PDSA cycles however, and when tested these ideas had only an average of about two PDSA run on them[7]
c	*Dosage/Coverage*		1 Collaborative, QI, 3 MDAs, Coaching and mentoring visits[17].	Only 44% of the respondents from the CEmOC survey reported that Neonatal Death Reviews (NDR) were being undertaken[7].

	Intervention Periods	Early intervention period (October 2008-September 2009)			
	Years	2008	2009		
	Quarters	Q4	Q1	Q2	Q3
E	Implementation process				
3	Execution				
a	*Fidelity*			Lack of resources at the end of the last financial year affected the timing of the Learning Sessions. As a result, in this quarter, sessions were done back to back to catch up with lost time, and it was necessary to combine the closing of 2008 Collaboratives and beginning for 2009 Collaboratives which was not ideal[22].	Clarification of the Salima Proposal focused on packaging what has been learnt and the essential elements necessary for spread. Proposal provides a starting framework for long term planning i.e. for spread and sustainability [24].
b	*Intensity*		Concerted coaching and mentoring visits are beginning to yield results in health centres as health centre staff take increasing responsibility for testing change ideas: Egs form Santhe, Kawaba and Kalulama HCs[20].	During HC Collaboratives: Challenges and progress for the learning period were shared, aims for the next stage of the collaborative agreed, a reminder of PDSA (plan, do, study, act) theory was made as well as clinic based PDSA, Process map review, change package introduction and selection of change idea and sharing of plans was done[23]. In Kasungu, there is now improved focus on the district hospital and its health centres centred on strengthening leadership, formation of a new QI team and strengthening of Kangaroo Mother Care[23].	Health Centre and CEmOC Collaboratives have changed the mindset of many teams on the provision of quality care for the better as they continuously learn from each other and develop new change ideas for testing[25].
c	*Dosage/Coverage*			Mentoring & coaching visits: CEmOCs were visited every two weeks, and most of the health centres every month[22]. 3 Health Centre Collaboratives in Salima and Kaungu and Lilongwe and 1 CEmOC Collaborative[22,23].	In Lilongwe, 110/145 planned visits were made by FI Officers. There has been an improvement in the visits of the FI Officers to facilities. Almost 80% of the visits are taking place; the 100% target was not met because of the Learning Sessions and attention being diverted to other needs, especially data deep dive follow-ups. The visits generally focused support towards QI Teams identification of high risk mothers and follow up, emergency referral checklist usage, and identification of blood donors[9]

	Intervention Periods	Intensified intervention period (October 2009-December 2010)				
	Years	2009	2010			
	Quarters	Q4	Q1	Q2	Q3	Q4
E	**Implementation process**					
3	Execution					
a	*Fidelity*			Methodology doesn't change over time but the areas for improvement and therefore change ideas that are tested do[9].	The design of the RCT is a constraint on the spread of the project[30].	
b	*Intensity*		CEmOC Collaborative: On the neonatal the teams will be working on reducing the deaths due to prematurity by having working kangaroo facilities with trained Kangaroo Mother Care Maids running these facilities[28]. In Salima, for the first time, the filling and sending of referral feedbacks to health centers came up to 100% and has been maintained since March'10. The team has now started doing resuscitation drills on birth asphyxia[28].		Rules of QI meetings: Data has to be presented at each meeting (FI officers to check whether data is being presented). Data was placed at the centre of the intervention. Data clerks were part of the QI team especially in the CEmOC facilities. Learning sessions were conducted with data clerks. Data sessions were conducted with super-improvers. However, the data was not very 'encouraging' since improvements in data collection and reporting usually showed an increase in the no. of deaths[9].	
c	*Dosage/Coverage*	In Kasungu, 40 visits out of 54 planned coaching visits were done to the health centers in this quarter, representing 74%[28]. This quarter has shown a drop in planned versus actual coaching visits and supervisions (to only 50% in two districts in January). The main reasons behind the drop were Christmas holidays, annual leave and the severe national fuel shortages in November and December[26]	Frequency of visits (dosage) only achieved 6 months ago; Using the DHMTs for spread, building on existing relationship between DHMT and HC Staff and their monthly visits[10].			

6.3 Intervention Characteristics

In this section I try and synthesize the following aspects of the intervention such as how well the intervention was designed, its adaptability and flexibility (i.e changes to the design over time), its ease of understanding (conceptual clarity) across different levels of stakeholders and its embeddedness within the health system.

6.3.1 Conceptual clarity

The original proposal, which was prescriptive of an integrated model for improvement, led to parallel work and very heavy reliance on numerous training sessions, which was a strain on the already stretched human resources within MaiKhanda. According to IHI, MaiKhanda's QI interventions was supposed to be based on the IHI collaborative learning model but had morphed into an untested model that was neither integrated nor scientific and unlikely to get sustainable results [XI]

Majority of facility staff claimed to have a broad understanding of quality improvement. The most frequent response for their understanding of QI was that it was the improvement of patient care or the testing of changes.

Some respondents who did report the link between CBA and PDSAs, reported as:

"*It is an objective, systematic and critical analysis of the quality of obstetric care against set criteria of best practice. This is by plan, do, study and act"* (QI Team Member Mlale Hospital).

They understood the integrated model as a continuous process. Figure 21 below illustrates how the integrated model was perceived by the facilities. It represents the integrated model as a complimentary set of processes that feed into the PDSA cycle. The PDSA test and implement ideas generated from the MDR, CBA and Change Package. In addition, there were recommendations that were developed directly from the MDR. The standards

fed into the CBA and also into the change package. The standards were considered important for setting how things should be in the facilities.

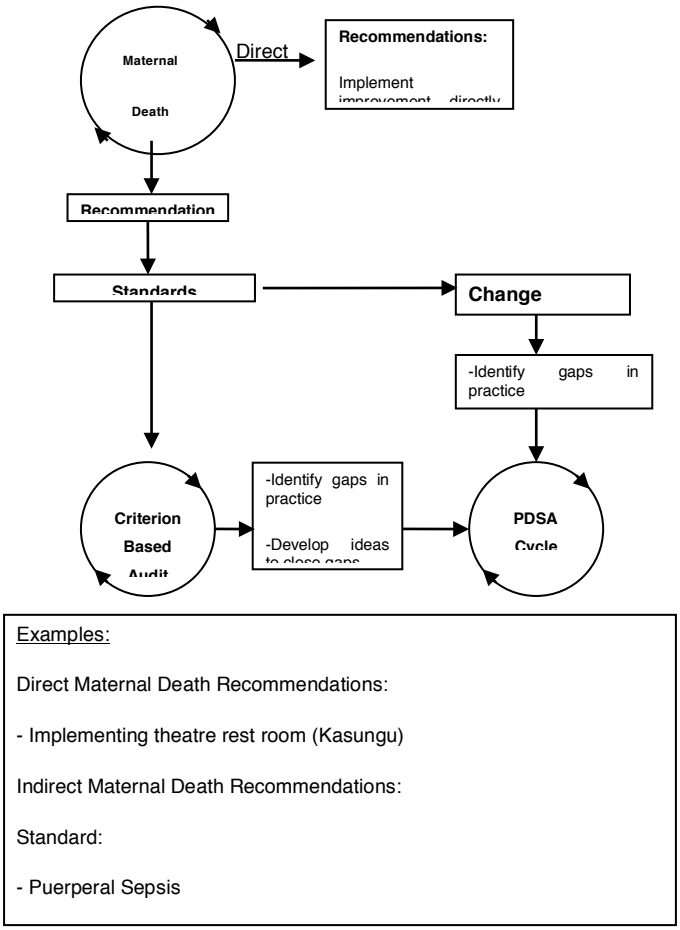

Figure 21: Facilities' perception of the integrated improvement model

The QI teams found their own way to use and link the QI tools-they had not necessarily copied the change package ideas identified by MaiKhanda. A number of innovative approaches to QI had been developed by the facilities. The teams were generating ideas independent to the combined approach by learning from other facilities.

However the MaiKhanda team was upset and confused by the different and disjointed approaches, particularly the links and overlap between criterion-based audit (CBA) part of the LSTM-led death reviews, and Plan-Do-Study-Act (PDSA) part of the IHI-led QI initiatives. There were uncertainties

expressed with regard to the effectiveness of the integrated QI Model [XIV]. As noted by a senior leader at IHI *"… they (the integrated model) have failed to work as effectively as needed in a coordinated manner, threatening sustainable progress and increasing the burden on local programme staff and participants"* [XIX].

This confusion persisted into the pre-intervention period as well with acknowledgement from THF (The Health Foundation), the donor agency that *"in the two approaches advocated by the IHI and LSTM, we (MaiKhanda) have been working with two very different paradigms for quality improvement"*[XIV]. IHI with support from the donor agency suggested that MaiKhanda embrace the IHI Model for Improvement as a unified, core technical framework for managing change. It was asserted that a general adherence to this approach would limit confusion for program participants and staff as they seek to transform their outcomes.

As a foundation to the early intervention period, the strategic review of programme intervention suggested "*A single collaborative improvement project as a unified social action system to guide improvement activities. A focus on keeping care providers in their facilities and deeply valuing subject matter expertise i.e abundant skill, good will, knowledge and energy of the local health workforce and the communities in which they reside'*[XIV]. This modification to the original concept meant revision of the intervention components.

Following this there was a directive from the MaiKhanda Director, in the early intervention period (2009), suggesting a move from "quality" improvement (QI) to "facilities" intervention (FI)[XX]. This reflected the intent mentioned in the above statement to consider QI as the overarching approach throughout the programme. In 2009, there was a move to describe the entire set of interventions (including the community interventions) within MaiKhanda as QI. Thus the quality improvement work in the facilities was to be re-worded as 'Facilities intervention (FI)' with QI being the overarching programme strategy for all of MaiKhanda's interventions. Discussion on this was triggered by

concerns among staff, especially at the district level, about MaiKhanda's "multiple identities" – something that seems to be perceived by the different groups of staff as well as by the stakeholders.

"It was observed that calling one component of our work "Quality Improvement" suggests that attention to quality improvement is limited to just that one part of the program. Hence, the (MaiKhanda) team agreed to change the terminology and that from now all our work must reflect a quality improvement approach, i.e. we must infuse a QI approach throughout the program (community intervention, management support services, monitoring and evaluation)."[XX]

This shift in the programme strategy from a facility based quality improvement intervention to having quality improvement as the overall strategy of the organization meant using the QI principles in the community intervention as well. This was never tested before and was not part of the original intervention design. In operational terms it was perceived as *"getting women to give birth at the facilities and then ensuring that they receive a 'skilled delivery' once they get to the facility"* as the primary objective of the project[XXIX].

External agents had a big role to play in developing the conceptual clarity around the intervention, with the donors subscribing to the changes suggested at different time points in the project (in fact, they were the ones suggesting and supporting this). However, in 2010, the donor reverted back to saying that "… *all partners have a role to play in their respective expertise areas and thus women's group interventions (though modified) and QI (also known as FI) were to be considered as independent interventions"*[IX]. There was also general agreement that getting women to deliver at the facilities was not the primary objective of the project [IX].

Some other general areas of conceptual dissonance are presented. While it was agreed that improvement would be implemented across a vertical slice of the health system, the unit of analysis were the health centres-this was discussed and agreed to in the proposal. Much of the work though, was focused in the CEmOC facilities initially. But this was justified since it was the

baseline period where the interventions were being 'tested' in the Malawian context. However work in the CEmOCs continued through the intervention period, with the rationale given by the implementing team that these facilities bear the maximum burden of deaths in the facility. The implicit assumption behind working along a vertical slice of the health systems and having health centres as the unit of analysis was that improvement of care in the health centres would help decongest the CEmOC facilities and thereby influence the mortality rate in these facilities. Secondly there was a greater emphasis on maternal mortality reduction and maternal change packages with newborn change packages being considered seriously only in the late intervention period. Thirdly, the original proposal was focusing on both mortality and morbidity-though morbidity was never addressed at anytime during the programme intervention-the focus being solely on mortality.

And finally, throughout the length of implementation, there was greater focus on improving outcomes rather than improving processes of care. Based on the literature for quality improvement, QI is focused on improving the process and this is done using tools such as statistical process control chart and run charts. In a health system context, it means improving the processes of care. The process of care itself can be categorized into organizational service delivery and the actual process of care giving, which in turn is based on the values and beliefs of the care provider and their relationship and position within the organization and with their patients. Therefore, in order to improve quality of care, QI interventions need to focus on improving organizational service delivery and influencing individual behaviour change along with improvement methods, such as PDSA, at clinical microsystems level. But in practice, the focus at MaiKhanda was largely on improving the outcomes (case fatality rates). Focus on outcomes, especially distal outcome, can be misleading and demoralizing for staff, especially if one looks at them at frequent time intervals.

In the late intervention period, there was a focus on improving the processes of care such as measuring and improving staff skills in neonatal resuscitation

within the “golden minute”[XXVII]. There was also a greater focus on engaging with the DHMTs [XXIV].

6.3.2 Intervention design

There were two models of improvement at the outset, which was supposed to be developed into an integrated model. These approaches, along with the Death Reviews and Data Improvement formed the core components of the QI implementation programme. But the recommendations from the original proposal[I] to develop an integrated improvement model never happened given the difference in approaches to QI amongst the two models.

The integrated approach as used by the consortium was specialized in that, criterion based audit identified the WHAT question-i.e the gaps in MNH while PDSA tried to resolve the HOW question ie how these gaps could be filled. Thus the rapid PDSA cycles would come into play as soon as facilities had identified their own problems and begun the analysis of possible solutions.

The death review (Audit) programme was based on MoH recommendations from the Malawi MNCH Roadmap. Death reviews was an established practice within the Malawian health system and MaiKhanda’s role was to facilitate the local staff to conduct audits in a blame-free and participatory manner. The advantage with integrating death reviews within the integrated improvement model was that it was a directive from the MoH. As MoH had identified MDRs as one of their priority areas, it was less likely to face resistance from key stakeholders, especially at the district and sub-district level.

A mid tem strategic review was held in 2008, where discussions were held with partners to see if there would be better ways of implementing the model[XV]. The intervention re-design suggested was the “*IHI Model for Improvement as a unified, core technical framework for managing change”.* Thus the IHI Model for Improvement was to provide the core framework for improvement activities undertaken by participating facilities.

Standards of care and criterion based death audits, which formed the basis for initiating improvement in the facilities, in the baseline and pre-intervention period, were not part of the intervention design in this revised model for improvement. The rationale was that the integrated model was incomplete as it included only one aspect of IHI's suite of tools, positioned within the context of CBA and thus focused more on expert guidance for individual health care providers rather than motivating teams and building their capacity to identify implementation issues in the local context, recognize system weaknesses and work as a team to test solutions (for improvement) [XIX].

The donor expressed their rationale to follow IHI's re-designed Model for Improvement and this culminated in the departure of LSTM (and subsequently LATH-the management arm of LSTM providing administrative support to the MaiKhanda programme) out of the Consortium.

By the start of the early intervention period, there was a strategic shift from an integrated MaiKhanda specific improvement model to an IHI model for improvement.

The revised activities included strengthened technical support to MaiKhanda and QI Teams, adherence to IHI model for improvement in health facilities, clinical skill building, data Improvement and leadership development, coaching and mentoring and technical support from MaiKhanda team. during action period.

Improvement activities were initiated from the change packages developed as part of the intervention rather than from death reviews or audits. These change packages were developed initially on the basis the death reviews but later on was developed based on 'gut-feeling' and data emerging from the facilities.

The intervention design, consisting of the model for improvement and breakthrough collaborative series remained stable throughout the rest of the implementation phase. The Rapid Cycle Improvement (PDSA) approach formed the crux of the intervention. It built on the providers' knowledge base

and was to be done through a structured series of small interventions, designed by the providers themselves, thus using a 'bottom-up', participatory approach. MaiKhanda provided a facilitatory and supportive role rather than direct implementation of the intervention..

Overall neonatal mortality rate in the MaiKhanda intervention areas had not declined in the 30 months when the intervention was active, partly because of less focus on this as compared to maternal outcomes. The implementation review team acknowledged that effective change packages were under-developed and there had been a lesser focus on neonatal outcomes [XV].

The revised improvement model was not designed with decentralised implementation for the facilities intervention-it assumed a more organic spread of the interventions from the CEmOC facilities to the health centres. However, the original model was to use a vertical slice approach, which would take into consideration the district health services structure. Moreover, at the time of commissioning the study, the health centres were chosen as the unit of analysis. Thus the focus of the implementing partner was on improving service delivery at the CEmOC facilities.

IHI had initially proposed a lean management structure (with only one QI officer to cover the 3 districts) with the assumption that the intensity of coaching would reduce over time. But, in practice coaching had to be on-going given the constant turn-over of staff and the fact that the concept of QI was new to the facilities. . The dynamic nature of staff movement within the facilities was anticipated but was not factored in the initial intervention design. It was clear throughout the course of programme implementation that stability of QI team was likely to have an effect on success of QI interventions, since it contributed to organizational memory and skills retention within the facilities. The initial intervention design had overestimated the capacity and the capability available within facilities to support the interventions. Some of the concepts also lacked clarity. For instance, super-improvers were trained for the CEmOC facilities but their position, role and status was not well justified in the programme. Furthermore, there was underestimation of human and

material requirements and logistics required by MaiKhanda to deliver the intervention. Data management strategy (ie linking the monitoring data from the implementation team with evaluation data from the evaluation team) was not designed from the outset.

The project was originally designed to be scaled up to all areas of the three districts within five years. This scale-up phase was postponed because the partners felt the current interventions had not yet been successfully implemented [XXIX].

6.3.3 Adaptability

Adaptability of an intervention is key to its success and sustainability. Adaptability refers to the degree to which an intervention is flexible to the local context. Although there was a technical working group within the Ministry of health in Malawi for improving quality, the concept of 'quality improvement' was fairly new in the country and adaptability of a 'quality improvement' model was often overshadowed by the 'quality assurance' approach propagated by the Ministry's technical working group on quality [XVIII].

Taking this into consideration, the original program design had proposed death reviews as part of its core interventions. In Kasungu and Salima districts, monthly reviews of (maternal) deaths that occurred in the district, were conducted at the district hospital while in Lilongwe, reviews were conducted during monthly review meetings, which brought together representatives of the various health facilities to the district health office. Standard Safe Motherhood (SM) treatment protocols for the major emergency obstetric complications, which were developed by the national safe motherhood project, were available in nearly all maternities in the three districts [II]. At rural health facilities, facility-based reviews were not done in any of the three districts.

Using the integrated QI approach (standards of care & death reviews), it would have been easier for MaiKhanda to create greater in-roads into the health centres as these interventions (i.e death reviews) were based along

MoH recommendations and therefore had better buy-in from the facilities [XII]. While death reviews were part of the original proposed integrated improvement model for MaiKhanda, it was not considered as part of the intervention after MaiKhanda's strategic programme review in 2008 [XVII].

The intensity of support provided by MaiKhanda's facilitation team was an important factor affecting the adaptation of the PDSA cycles in the health facilities. In the initial stages of the intervention, the programme was implemented largely by the technical experts from the Consortium without much engagement of the MaiKhanda staff [X]. The idea was to hand over the knowledge and expertise directly to the QI teams on the ground. By the early intervention period, MaiKhanda was established as a local NGO within Malawi (rather than a project) and programme implementation through MaiKhanda was encouraged, with the Consortium members providing technical support where required [XV]. However, in the early intervention period there was only one QI officer responsible for all three districts [XIV], and this did not change until 2009 (the late intervention period) where there were QI officers at the district level and also for the CEmOC facilities [XVIII].

Adaptations depend on the frequent testing of ideas. Conceptually, the PDSAs are rapid cycles of intervention done on as frequent basis as possible. There were assumptions regarding the 'rapidity' of the action cycles. In reality, the rapid action cycles could not be carried out at the speed and frequency, which was expected of the intervention. From the CEmOC survey, for the period form June 2006 to August 2008 a total of 92 QI activities (ideas to test) were reported to be conducted amongst the different CEmOC facilities. Of the 92 ideas that were generated through the change packages, 26 ideas did not have any PDSAs run on them. For the 66 ideas that were tested, there were about 180 PDSAs run, giving an average of 2-3 PDSAs for each idea to test [VII]. This ability of QI teams to adapt the PDSAs to their local context was related to factors such as stability of the QI teams in the health facilities and their grasp of QI concepts. Although coaching was to be provided through an

initial intense phase followed by phasing out, in reality, coaching continued throughout the length of the project, given the staff turnover in the facilities.

Adaptability of the intervention was also limited by the length of its implementation chain. The MaiKhanda interventions were characterized by long implementation chains. The programme strategy was handed down from the technical lead of consortium partners to the local team (MaiKhanda) on the ground. The MaiKhanda staff in turn built capacity of the implementation teams (QI teams) in the health facilities and facilitated the intervention. The QI teams were ultimately responsible for implementing the intervention on the ground.

6.3.4 Embeddedness

Leadership buddy system, developing local leadership, QI champions, exchange visits were all additional activities planned in the original proposal for embedding QI into the health system [I]. MaiKhanda's interventions aimed to work alongside the Malawian Ministry of Health (MoH), in achieving their national 'Road Map' for reducing maternal and infant mortality. This included developing & implementing appropriate behaviour change interventions conduct maternal death reviews and clinical audit & strengthen health facilities at all levels with functional EmOC services [XII]. Maternal death reviews in health facilities, in line with Ministry of Health needs were used to identify problems and implement solutions in the baseline period of the intervention [XII]. However, facilities did report encountering resistance for their QI activities from other organisations working in their facility, during this period.

"… their newly developed labour graph met resistance in the sense that the Safe Motherhood Initiative authorities said we should hold using it"(nursing staff, KCH) [III]

While there was a strategy is to work closely with both health centres and DHMT (District Health Management Team) so that critical needs could be incorporated and budgeted for in the district Implementation Plan [XIV], MaiKhanda only engaged in strategic MoH meetings towards the end of the pre-intervention period [XX].

There was engagement with RHU and DHOs in MaiKhanda's strategic planning meeting which formed the basis for the final intervention. Capacity strengthening of MoH staff on data collection, interpretation and dissemination was done so that routine health facility data could be of good enough quality to be used for improvement purposes by the local QI teams. [XX].

Towards the end of the late intervention period, it was noted that many of the breakthroughs in “women friendly care” and in standardizing data collection were being taken for granted by the facilities and new members of MaiKhanda team [XVII], suggestive of a degree of embeddedness of the women friendly change package within the health care system.

Also, during this period, the MaiKhanda’s District Coordinators were to be entrusted with a central role in working with the Ministry of Health officials, to help get key changes developed and tested during facility Collaboratives, to be embedded within the district health care systems. This would include changes in MaiKhanda’s policy, guidelines, forms, job descriptions, and on-going training of health care workers. A bigger roles for District Coordinators was envisaged as the QI intervention moved from testing to embedding within facilities [XVII].

However, toward the end of the late intervention period, it was documented that most changes were not being reliably implemented and had not been fully integrated into local, district, or national structures of the Malawian health system. This was attributed partly to the randomization design [IX, XXX].

6.4 Health Systems Context

Context can be considered at different levels like the rings of an onion, but unlike the onion rings, they do not have strict boundaries and there is usually an iterative and dynamic interaction between these different levels of the context. The different levels that I have considered here include the external environmental, the health systems context, the organizational context within the health facilities and MaiKhanda’s own internal context.

6.4.1 External environment

The level of political commitment for maternal and newborn heath in Malawi was very high as evidenced by the commitment to the Road Map proposed by the MoH. The Road Map for maternal and newborn health aimed to develop and implement appropriate behaviour change intervention and advocacy strategies to address maternal and newborn health issues, to conduct maternal death reviews and clinical audit and strengthen health facilities at all levels to ensure adequate coverage with functional CEmOC and BEmOC facilities.

At the time of initiation of the project, the political environment in Malawi had been quite stable over the past year. With the government bid to end corruption there was confidence in money being used optimally thereby providing greater confidence among donor agencies to invest in the health sector in Malawi. A good example was that of SWAp (Sector Wide Approach), the EHP (Essential Health Package) and EHRP (Emergency Human Resource Package) being implemented across the health sector with commitment from different stakeholders. The Emergency Human Resources in Health initiative supported by providing salary top-ups and other incentives for health care workers(210).

With Malawi receiving funds from GFATM (Global Funds for AIDS, Tuberculosis and Malaria) for health systems strengthening, 'global HSAs' were trained and deployed by the government with each HSA covering a catchment population of not more that 1,000 (while previously it was a population of 4,000). These initiatives although directed towards broader health systems, had an indirect influence on maternal and newborn health in the country(10, 12, 214).

The quality improvement programme facilitated by MaiKhanda came into Malawi at a time when the maternal and newborn mortality in Malawi was comparatively very high (§1.2). Quality was beginning to be considered an essential element of care and various interventions such as the quality assurance project by JHPIEGO was being institutionalized in many district

hospitals. Ministry of Health had also constituted a Quality Assurance Technical Working Group.

Despite these national level efforts, there were challenges in the facilities. Implementation of essential BEmOC signal functions, was erratic and most facilities did not qualify as a BEmOC site. The MaiKhanda programme had *effective supportive systems* as one of its change packages. Death reviews as supported by MoH had little value as it was mainly for reporting purposes with hardly any action being taken on the recommendations. MaiKhanda through its innovative tools and techniques such as rapid improvement cycles (PDSA) and criterion based audits hoped to support QI teams in making the recommendations from the death reviews actionable. The establishment of change packages was also aimed at bringing about improvements in maternal and newborn care in the health facilities at the microsystems level.

Despite all the external initiatives characterized by donor funding and external budget support, availability and utilization of drugs and supplies at the district and CEmOC facilities was perennially inadequate [II,IV,VII]. From the facility resources survey, it is clear that districts with better resources performed better, despite staff shortages. For example, Salima district was a role model district, with its better availability of drugs and materials was relatively better performing district as compared to Kasungu and Lilongwe [XXI, XXV].

With the EHRP, the donor community was trying to address the gap in human resources existing in Malawi. However, the problem of human resources availability within the health system persisted throughout the length of MaiKhanda project and this had implications on how the interventions were delivered on the ground. The problem persisted through the early and late intervention period and it varied across the districts [IV]. The health facility resources survey of 29 HCs showed only 14% of the staff were clinical, while HSAs constituted the bulk of remaining non-clinical staff. Of the clinical staff, majority of them were nurses (71%).

At the time of the baseline survey in 2006, it was a national policy that at the ante-natal clinic women were requested to bring with them when they come to a health facilities to give birth, a clean plastic sheet to lie on, a clean razor blade to cut the cord, and two chitenjes *(pieces of cloth worn as a sarong, usually by women)*, one to cover herself and the other to wrap the baby. Most women who delivered in health facilities would adhere to this request [II]. It is not clear whether this policy had an effect on the perceived quality of delivery care at health facilities. However, this policy was no longer in existence when MaiKhanda started its project in Malawi in 2007.

Another policy of the government at that time, which affected the intervention in the early intervention phase (2008), was the comprehensive ban on traditional birth attendants by the government to conduct deliveries at home(215). This was supported by village chiefs and Traditional Authorities, who imposed ban (such as a goat or chicken) for women delivering outside of facilities. This had implications for programme implementation as facility deliveries went up significantly between 2007 and 2010, from about 40% of deliveries to about 80% of deliveries (Figure 22). The ban got lifted in 2010 as quickly as it was imposed.

Programme implementation was also affected by competing programmes of other partner organizations. Other MNH projects that were being implemented in the health facilities included: Antenatal Care Improvement, Mother to Mother/, IMCI (Integrated Management of Childhood Illnesses), CBDA (Community Based Distribution Agent) supported by MSH (Management Sciences for health), Advanced Life Support Obstetric Course (ALSO), CESTA (an Italian NGO) which provided refresher training to midwives and allowances for locum shift midwives, home-based care by MSH, reproductive Health sponsored by UNFPA, Infection prevention, PMTCT Programme [VII]. DHOs in both Kasungu and Lilongwe were also working with other partners like MSH and JHPIEGO who were involved in quality assurance interventions and other health related programmes. As such sometimes it was difficult for the implementing partners (QI teams) to concentrate fully on MaiKhanda programme activities. As much as

MaiKhanda was competing for the same pool of local talent as other partners, there was also competition in terms of ideologies regarding quality improvement.

Fuel prices had gone up impacting on programme budgets [XVII]. The global rise in fuel prices has also had a negative impact on transport for project implementation.

6.4.2 Health systems readiness

6.4.2.1 BEmOC and CEmOC facilities:

At the time of initiation of the MaiKhanda project in 2006, Malawi had almost double the recommended minimum number of Comprehensive EmOC facilities (1.8 facilities per 500,000 population) but only 2% of the recommended number of Basic EmOC facilities (0.1 facilities per 500,000 population). There is no equity with respect to geographical distribution. The met need for EmOC was about 18.5%, which is far below the UN recommended level of 100%. Quality of EmOC services is generally poor as seen by a case fatality rate of 3.4% which is much higher than the UN recommended level of less than 1% (216). In the baseline survey conducted in 2006, none of the surveyed health centres could be classified as a fully functioning BEmOC facility. None of them had all the 6 signal functions working at the time of the survey [II]. This was similar to the findings in the EMoC assessment survey conducted by the RHU.

Maternal and newborn health services were provided by 13 hospitals and 60 health centres in the three districts where MaiKhanda was to start its project implementation. Nine hospitals functioned as CEmOC facilities - one in Kasungu, one in Salima and seven in Lilongwe. None of the health centres could provide all the BEmOC signal functions at the time of the survey. LLW has a population that is nearly 6 times the size of a normal district. The HCs have on average 10,000 additional population to deal with thereby increasing workload [VI].

6.4.2.2 Care provision

In the early part of the project, all three districts of Kasungu, Lilongwe and Salima, the proportion of expected deliveries conducted in health facilities, was around 40%. There was however, district level variations with the proportion of expected deliveries being higher in Lilongwe (45.4%) as compared to Kasungu (27.7%) or Salima (28.9%) [II]. In general, there seemed low utilization of delivery services. Only a limited number of signal functions were available at the facilities and that too remained inconsistent(33).

Almost all health facilities were baby-friendly facilities- babies were wiped dry and wrapped in a cloth after birth, given to the mother shortly after delivery and the mother was encouraged to initiate breastfeeding the baby. Screens or curtains were available in the delivery room in 59 out of 73 health facilities (81%). In most health facilities women are allowed to take food and move around while in labour. However a companion – such as her mother or sister - was allowed to stay with the mother in the labour ward only in 22 out of 73 health facilities (30.1%). A bed-sheet to cover the mother while in the delivery room was available in 26 out of 73 maternities (35.6%) [II].

In the pre and early intervention period, while facility QI teams and other health workers were keen on improving the quality of care in the early intervention period, in many facilities (e.g. in Mlale and Bwaila hospital) overcrowding in nursery and delivery areas and shortage of equipment and supplies (e.g. laboratory reagents) were major constraints [XIX].

For Lilongwe, the more urban a health centre was, the more nurses it had and thus also lower nurse to delivery ratio. This perhaps suggests an inclination of health centre staff for work in more urban areas with their amenities and access to better schools for their children. Despite this staff motivation is lowest in Lilongwe as compared to other districts [VI]. In Kasungu and Salima, health centres are generally more rural. However, Salima health centres appear less disadvantaged than those in Lilongwe and Kasungu, with more of

them having access to electricity, mains water and tar roads, lower nurse to delivery ratios and lower case fatality rates [IV].

6.4.2.3 Referral system

Long distances to health facilities, particularly emergency obstetric care facilities, and lack of means of transport were major problems affecting the referral system in the rural areas of the three districts, while during rainy season navigating the un-tarred roads was difficult. At baseline, Salima and Kasungu districts had a better-organised referral system, based on a functional radio communication system between health facilities, so that they were able to call for an ambulance from the district hospital when an emergency case needed referral. In Kasungu, all health facilities had a radio communication system or a ground telephone line, and a car ambulance available in 5 out of its 28 health facilities (28%). In Salima, all health facilities had a functional radio communication system, but only two health facilities had a car ambulance. In Lilongwe all health facilities had radios, but only 13 out of the 42 (31%) were functioning, and 13 out of 42 (31%) health facilities had a car ambulance [II].

6.4.2.4 Death reviews

At the time of baseline survey in 2006, few maternal deaths (MDs) happened at rural HCs and most MDs occur either in tertiary hospitals or at home. In Kasungu and Salima, regular reviews of maternal deaths (MDs) that occurred in each district was conducted at the district hospital on a monthly basis. In Lilongwe, MD reviews were conducted at monthly review meetings, which brought together representatives of the various health facilities to the district health office. At rural health facilities, facility-based MD reviews were not done in any of the three districts. While, there was an expectation from the MoH that health facilities would complete all their own MDRs, in reality the facilities had many challenges. For example, Kamuzu Central Hospital (KCH) experienced a large number of maternal deaths, and often there were not enough personnel available to complete each MDR.

Peri-natal or neonatal deaths reviews were reported to be done at a few health facilities only [II].

The focus throughout the initial part of MaiKhanda project was largely on maternal death review. Neonatal death review had not been initiated in most facilities at the time of the survey. The focus on reducing neonatal deaths was not as strong as maternal deaths because facilities chose to focus on maternal deaths. This is because health centres did not have the equipment such as ambu bags that would help them with managing neonatal complications. But even with equipments, there was still need to create a system in place before hand[XXXI].

6.4.2.5 Resource availability:

A serious set-back to the provision of quality maternal health care services was the critical shortage of qualified health workers, especially professional midwives. Most health facilities in the three districts, particularly health centres in remote rural areas, were understaffed and health centres usually were staffed with only 1 or 2 midwives, which was inadequate to ensure continuity of maternity services. In some rural health facilities the nurse-midwife also had attend to general outpatients as well, which also had an affect on the quality of maternal health services being provided [II].

Shortage of staff meant that staff transfers lead to void in the facilities where the staff were moving out from. In addition, in the bigger CEmOC facilities there was also the issue of being transferred within the hospital to other wards. Thus, health workers that were trained in QI either left or were transferred thereby creating a gap in the continuity of the intervention [VII]. In theory, training and coaching in QI at MaiKhanda was to be done in a phased manner, such that local champions were able to take lead in QI coaching over a period of time. But in reality the high turnover of staff meant that coaching had to be continued throughout the period of the project and this slowed down the quality and progress of QI activities in most facilities that MaiKhanda was working in.

Another issue related to human resources was the non-availability of the already limited staff in their facilities given the vast number of training and workshops that were being conducted by various stakeholders in the country.

Despite all these trainings, there was a critical shortage of neonatal clinicians and neonatal care skills in some facilities this led to reduced attention to newborn care. For example, Kasungu did not have a paediatrician for 10 years and when provided with support for QI work, they requested a paediatrician for 2 days a week [XIX].

Cleanliness of the maternity environment, availability of water, electricity for lighting, equipment, drugs and supplies have a bearing on the quality of maternity care. Several rural health centres relied for their water supply on borehole water, which was kept in buckets in the labour ward [II]. The health facility resources survey conducted later in the programme was not any more encouraging than the findings that were at the baseline period. While (52%) of the health centres have both electricity and water supply, almost one third of the health centres (31%) had neither water nor electricity. These were mainly the rural health centres, which also had poor roads [IV]. However, most of the health centres in Salima were also rural, but more than half of them had road access and all of them had access to water supply and electricity.

Health centres are very under-resourced in general with the minority having functional ambulances, some not even having functional clocks in their labour ward or functional BP apparatus and most not having clinical officers available at night and some not even having nurses available at night. While 14 (48%) of health centres had all the materials available, only 9 (31%) health centres had all the materials functional.

In terms of comparisons between districts there were some differences such as Salima having fewer health centres with functional ambu bags but more with functional car ambulances, and more with functional clocks in their labour wards; Lilongwe having fewer health centres with functional radio communication. Salima health centres are generally better off than Kasungu

and Lilongwe health centres with respect to electricity, water and deliveries per nurses per year; but they were worse off on indicators such as deliveries per bed.

At health centres supplies of oxytocins and antibiotics had not improved from the time since the interventions were initiated- half were without these essential drugs at any one time. Most of the 62 health centres provided only 1 or 2 of the 6 signal functions of basic obstetric care and some not providing any at all. Lilongwe health centres, with an average of 1.0 of the six BEmOC signal functions available in any given health centre in any given month, were worse off than Kasungu health centres (average of 2.3) which are worse off than Salima health centres (average of 2.9) [IV].

6.4.2.6 Organizational Culture

Based on the findings from the Staff Psychology Survey and CEmOC survey, the organizational culture in Malawi is very hierarchical in nature [VI,VII]. The relationship of staff with their supervisors was perceived to be good, but consultations by the senior leadership or supervisors on work-related decisions was usually lacking. This lack of freedom to take decisions about their work, meant that it was more difficult for staff to test out new ideas using the PDSA cycle. An environment fostering innovation and autonomy by staff to carry out such improvement work was apparently lacking at times. Staff also complained of a lack of supervision by their supervisors [VI].

Nurses were worse off than medical assistants when it comes to relationships with other staff and colleagues in their setting [VI]. Medical assistants were less engaged with the patients as compared to the nurses. Urban health centre staff were worse off than their counterparts in rural areas in terms of their motivation, interest in the current job and relationship with colleagues and other members of staff in their setting.

The climate in the facilities was generally pleasant-staff got on well with each other, could discuss problems with colleagues, encourage each other and had good relations with their superiors.

Staff are not always accommodating to patients. A third of the respondents, from the staff psychology survey, said that staffs speak rudely at their health facility some of the time. Only 50% of the staff said that they would recommend a friend/relative to deliver at their facility [V]. Government staffs were consistently worse than CHAM staff facility in this regard. But there are differences between districts as well. For example, Salima appear on average to be more women-friendly than those in Lilongwe [VIII]. This resonates with other studies, which indicate the existence of a relatively better health service in Salima [IV]. In the intensified intervention period, 37% of the respondents (mainly nurses) felt that women should labour in the position of their choice [V]. There are other differences as well between the districts on a variety of determinants. LLW is worse of on issues related to happiness, job interest and conflict with patients [VI].

In general, CHAM facilities (for example Likuni Mission hospital and St.Gabriel's) were more active in implementing the change packages using the improvement cycle as compared to government facilities.

While the social network within the facilities were reported to be good, the network between facilities was poor [VI]. The QI Collaborative provided an opportunity for links between health centres thus improving the potential for spreading QI work between facilities.

6.5 Internal context within MaiKhanda

This section describes the evolution of MaiKhanda as an organization and includes its structural characteristics, governance and oversight mechanism, donor role and partnership, network and communication and capacity within MaiKhanda.

6.5.1 Structural Characteristics

To begin with, although the Programme had 37 staff members on board for the project, for the QI intervention, there was only one officer at the beginning of the project responsible for QI and another officer responsible for criterion-based audits and maternal death reviews, both based at the central office in Lilongwe [XV]. MaiKhanda had established district offices, but these were not utilized for the purposes of QI intervention.

The reason for adapting such a centralized lean model was based on IHI's theory of improvement as well as their experience of having done similar work in Peru and certain regions in Russia. IHI's model for improvement is itself based on the approach of diffusion of innovation, where innovation and ideas for improving care surface in local settings and the IHI role was that of a catalyst to help cultivate these innovative concepts and facilitate the spread of these ideas through peer-to peer learning such as the "Breakthrough Collaborative Series". Their previous experiences in Peru and Tula Oblast region in Russia had used similar team based approaches and were hailed as success stories (126, 169).

The plan for the project in Malawi was to develop a focused change package and begin initial work in the Central hospital in Lilongwe and CEmOC facilities, providing support to the clinical leadership and identify 'champions'. By working with District Health Management Teams (DHMTs) and having Collaborative learning sessions across the 'vertical slice' of the health system, the project expected to cover two-thirds of all the facilities in the 3 districts over a period of 2 years. 'Knowledge Agents' were to be identified from within the Ministry of Health who would be trained and coached, so that they sustain the interventions in the long run[I].

The intervention design, therefore, envisaged a single QI officer operating centrally from Lilongwe facilitating activities in all 3 districts and in the CEmOC facilities. The focus was on district and central (CEmOC) hospitals while continuing to work with smaller hospitals at same time [XI]. The focus on health centres (the main unit of analysis for the cRCT) was less of a priority in the pre-intervention period, the rationale being that pre-intervention period was used mainly to understand the health system and also to test the combined model of improvement.

However, the rapidity and spread of the intervention was perhaps over-estimated. Towards the end of 2007, the Collaboratives were still focusing on the CEmOCs and mainly on maternal change packages. The rapidity of spread of intervention as originally envisaged for the QI intervention was also perhaps dampened by the consideration to develop a combined model of improvement, which would integrate the criterion-based audit with the rapid cycle improvement (PDSAs). This was complicated by the fact that part of the integrated model, the CBA and MDR, managed by the MNH Programme Officer (with its affiliations to LSTM), reported through MaiKhanda's existing management structures while the rapid cycle improvement component of the intervention was reporting directly to the Technical Leads, bypassing MaiKhanda's organizational structures [XIII].

This complicated structure was thus perceived to be incapable of producing a cohesive outcome, leading to waste and duplication on the ground, undermining the role of the in-country team and demoralizing staff [XIV].

A re-design of the intervention was approved by the Board of the donor agency and this had implications for the structural and functional aspects of the programme. The changes included the appointment of a full-time Programme Director within Malawi, who would be the single point of contact for the Consortia partners, a single integrated workplan with a single programme budget. In addition, a Faculty Lead from IHI would be the principle technical advisor to the Programme Director in order to support the

development of an integrated workplan covering all elements of the technical support [XIII].

In the early intervention period (i.e 2007-08), with the appointment of a Director the programme transitioned from a project to a locally registered NGO in Malawi and came to be known as MaiKhanda. Quality Improvement was proposed as the overarching strategy for the programme, with the quality improvement in the facilities being renamed as facilities intervention. This facility intervention had some new team members joining them namely a Data Coordinator and Asst. QI Officer. There were also plans to have district level QI Officers but they were not recruited until July 2009 [XXIII].

The logistics of engaging the relatively large number of health centres from across the 3 districts, mainly for the QI work, was not really well thought through in the early intervention period. The other components of the intervention such as the community intervention work and the evaluation team in the districts working through the District Co-ordinators. But, for QI, the Officer worked directly with the QI teams (both district and central), bypassing MaiKhanda's district structures.

There were constant changes to the QI staff at MaiKhanda during the later intervention period that affected the continuity in supporting front-line QI teams [IX].

6.5.2 Governance and oversight

There was discussion at the beginning of the project in 2006, in particular of the need for the workplan to reflect the attempt to create a single, shared model of QI and all activities to be owned and genuinely led by the team on the ground. There was also a plea for greater degree of clarity about visits to Malawi by overseas individuals, particularly roles of Technical Advisors (TA). It was agreed that all visits, their timing, purpose and agenda need to be driven by local needs, rather than by what suits the diaries or interests of TA providers [X].

There were misunderstandings concerning the integrated model, and with regards to which partner should take the lead in QI. This resulted in the perception that there was no ownership for the integrated model. There were also discrepancies in how MaiKhanda's consortium partners identified themselves when introducing themselves in the facilities. For example some staff members introduce themselves as MaiKhanda, whereas others affiliated themselves with their parent organization (such as IHI or LSTM). This inconsistency in the identification of MaiKhanda as one team resulted in confusion at the facility level [III].

During the pre-intervention period, the management capabilities of the team on the ground saw the programme evolve within a short span of time despite the initial delay in implementing the project. However the team felt that they were not being given adequate breathing space to develop their full potential given the differences of opinion among the senior members of the consortium on the way the programme was to be implemented. This caused a lot of grief and frustration among the forbearers of the programme on the ground affecting staff morality and level of functioning [III].

The interventions in the baseline period were implemented directly by a resident improvement advisor from IHI. Local management structure was not clear before the start of the intervention and evolved over a period of time.

As much as there was tension regarding the integrated model for improvement, it partly stemmed from individual partners perception regarding their position within the organization. IHI considered itself the lead partner in the organization while LATH was responsible for managing the activities of the Consortium in country.

During the baseline period, the team in Malawi was lead by a Programme Manager, but the extent to which some lines of accountability within the team were not clear. The Programme Manager was line managed by LATH on behalf of the whole Consortium. LSTH, IHI and WCF were to provide advice to the Programme Manager about the technical content of the intervention. The

QI officer was line manage by IHI directly, but as a member of the Malawi team, was accountable to the Programme Manager. The QI officer was an expatriate and a direct employee of IHI, but there were plans in place to identify a local counterpart who would eventually takeover the role. There was this implicit assumption that such a local QI officer would be under the direct line management of the Programme Manager. This was made explicit during the mid-term strategic review in 2008 [XVII].

The idea to promote local ownership gained prominence in 2008 characterized by the establishment and registration of MaiKhanda as a local NGO and the appointment of a Director for the programme. In the early intervention period, MaiKhanda Senior Management Team was formed which discussed and developed a "committee" structure of management which was shared and discussed with the full MaiKhanda staff during the annual strategic planning meeting in 2008 [XVII]. A three-tier management structure was adopted: (i) Senior Management Team (SMT) based centrally in Lilongwe; (ii) Operational Management Team (OMT) consisting of all the members of the SMT and the three District Coordinators; and (iii) the District Management Team (DMT) consisting of key MaiKhanda district staff [XX].

Since the establishment of MaiKhanda as a NGO under the local leadership, the Director reported a significant improvement in MaiKhanda team wellbeing "*with a growing sense of cohesion and capacity for collegial decision-making'*'[XX].

By the beginning of the early intervention period in 2009, governance was affected by the absence of quality improvement staff in the districts and the generally weak communications between district staff and the central office quality improvement team. This was seen as a constraint to MaiKhanda's district-focused decentralized approach. It was perceived by district level stakeholders as though there were two separate organizations in operation -- a community-based one and a facility-based one [XIX].

Despite the move towards local in-country leadership and a furthermore decentralized district based approach, technical advisors still played a major

say in the programme management leading to misunderstandings between the local team and the technical partners. Strained relations between MaiKhanda Director and IHI Technical Lead in the late intervention period were a further detriment to smooth governance on the ground. Concern were raised over who ultimately makes decisions. The Programme Director criticized the consortium governance stating that *"…the Salima Proposal and the Communications Strategy died because the decisions are ultimately made 'over there'"*. [XXXII]

6.5.3 Donor role and partner relationships

There were frequent strategic changes to the intervention during the lifetime of the project, mainly influenced by the donor agency and technical partners.

At the beginning of 2007, after the first combined workshop in February, 2007 there was support for the integrated model for improvement from the donor. There was a strong recommendation from the donor that "*the 'lead' column should be removed from the workplan and… for all activities to be owned and genuinely led by the team on the ground, not by overseas organisations.*" [XI] However by end 2007, there was recognition by the donor that the integrated model was not viable. It felt that the 'integrated' approach was not materialising in practice, in large part because there were separate budgets were being held for the individual strands of the work, separate workplans for the different technical inputs and separate decision-making processes, often with key decisions being made outside the country. In November 2007, The Health Foundation issued a statement aligning with IHI model for Improvement, calling for a reorganization of the set up. Thus the donor engagement in the project had a great influence on the intervention planning and design and eventual roll out of the interventions [XIII]. A pre-intervention period had to be considered, since the intervention was not considered ready for uptake, which lasted for 20 months [XIV].

LSTM and IHI were the main partners responsible for the QI intervention. Although the partners were responsible for different components of the QI work (LSTM – MDR, standards and CBA, IHI – PDSA, data improvement) as

MaiKhanda approach aimed at integration, so the roles of the partners would ideally also have been integrated. For example, IHI would use the information available from MDRs to facilitate the development of ideas, which would be implemented in the PDSA. The MDR and the CBA were used by LSTM for identifying problems, and areas for improvement, and then the role of IHI was to take these problems and work with the teams to find and test ideas to rectify these problems, before implementation. Therefore there needed to be a strong link between the work going on with MDR, CBAs and PDSAs.

However, the senior members of partner organizations did not necessarily agree with this shared sense of responsibility. They felt that a lot of resources and time had been spent working on the collaboration and coordination, and that the integrated approach had not worked successfully as originally envisaged and that the integrated model was not implemented effectively in the facilities. The language of partnership was felt as lacking between the different partner organizations and therefore, while MaiKhanda had a very strong and committed staff working on the ground, the affiliation of these staff members to their parent technical partners affected their functioning as a coherent team. The implications of such a disintegrated approach was more subtle. It was not reflected in the day-to-day activities performed interchangeably between the staff members rather, the implications were more intrinsic. Various team members during their interaction with the evaluation team for the baseline evaluation, had expressed increasing amount of frustration in performing their duties as a 'team' [XII]

"… *we need to get clarity in terms of what the combined approach is...we need to blend the expertise, and for us to blend we need to reach a consensus"* (MaiKhanda staff member) [XII].

The donor's dual role as 'partner' as well as 'donor' added to the complexity of the arrangement. There was a lack of clarity among the consortium members as to what the donor's role as partner was. In addition they made it clear that their affiliation was with IHI's model for improvement, having collaborated on previous projects with them in the NHS in UK [XIII].

However, by the early intervention period, The Health Foundation clarified that MaiKhanda was a decentralised program, and that IHI is the lead partner with a role in QI technical execution and other partners also have roles in executions of their technical areas [XXII].

6.5.4 Network and Communication

The programme design had outlined channels of communication amongst Consortium partners and with the MaiKhanda team. There were also regular forums for communication such as the biannual Programme Management Board (PMB) meeting (which included MaiKhanda facilitation team, technical advisors, consortium and donor representatives). In addition, monthly and quarterly progress reports were shared with all partners. Yet communication was identified as a challenge in the beginning of the programme [XV] and remained a challenge throughout the programme.

It was observed in the baseline evaluation (2007) that IHI was deviating from what was agreed upon in the proposal and more importantly, "*There have not been any consultations or even communications with the other partners" about these deviations".* There were a number issues being raised with regard to communication within the Consortium, for example appropriate staff members were not receiving the correct information [III].

There were other instances when partners expressed concern that they were not fully informed on what was going on in the Programme – i.e. they were not engaged in discussions regarding the strategic shift from quality improvement to facilities improvement and QI being the overall unifying programme strategy. Secondly, they had not been informed about what activities the Technical Lead had been doing [XV].

Initial challenges were with the external consortium partners and the MaiKhanda team. Partners continued to communicate directly with their responsible Program Officers without involving the Program Manager. This settled in 2008 with the appointment of a Programme Director. But as the

programme became decentralised there was a need for strengthening communication between the MaiKhanda central team and districts.

Nevertheless, internal communication improved in the early intervention period, with the Programme Director focused on understanding and reviewing the internal management systems and operations and working with staff to make improvements at two levels, namely the strategic management level and the operational level. Despite all the structures and forums, communication still remained a challenge between districts and with central office. Many problems were identified due to lack of communication. Stakeholders in the districts did not always see MaiKhanda as one project because MaiKhanda staffs from the districts were not involved in QI activities. *"Our friends from the central office just come into the district and we sometimes just hear of events happening via the partners. MaiKhanda central office needs to fully engage district staff in QI so that they know what is happening"(Kasungu District Office).* [XIX]

Another example was that on allowances – each district had been choosing its own levels of allowances and this was confusing to partners who thought of MaiKhanda as a set of different organizations. Communication from central office was also identified as being problematic, as one of the District Coordinators quoted, " *Many times you just get a call or an email saying things have now changed and you are to do this or do that without any written formal communication stating the policy; written communication is important. Written communication to District Co-ordinators needs to be streamlined*" [XIX].

Scheduling and maintaining appointments for follow up at HCs was often difficult, largely due to limited staff available at the district level. This was resolved with the appointment of district QI staff but then there were other issues such as lack of transportation.

Towards the end of the early intervention period, communication from the Technical Lead was mainly with the Facility Improvement (although it had been agreed that QI would be the overarching theme for MaiKhanda), by-

passing the MaiKhanda organizational structure leading to conflict with the role of Programme Director.

6.5.5 Capacity within MaiKhanda

During the pre-intervention period, MaiKhanda did not have officers responsible for QI activities at district level. With the rolling out of QI activities into health centres, it was extremely difficult for the available QI staff to conduct mentoring and coaching visits to all facilities and follow up on the teams projects. There was a felt need to consider having district based QI officers who could provide on-going support to QI teams at district level. Subsequently, a number of training activities for MaiKhanda was recommended and infact, also took place [XX]. These included training in life saving skills for MNH officers, improvement advisory training for QI programme, training in financial and accounting system. Others included super improvers (from facilities) training and establishment of an Improvement Advisory Board (although this was not sustained through the length of the project). Thus, there was a special focus on capacity building in the early intervention period. There was a significant focus and investment in professional development for both existing team members and new recruits. This included initial orientation of new FI members with support from IHI technical advisors, and an ongoing on-the-job coaching and mentoring of the entire team by IHI technical team [XXIII]. In addition, senior MaiKhanda staff were on the IA (Improvement Advisor) course supported by IHI. A draft training plan with areas of need was developed.

Recruitment of district-based officers had stabilized and improved focus on QI activities with more frequent coaching visits to QI teams, which was critical in influencing QI application on the ground [XV].

By early 2010, MaiKhanda staff had completed their Improvement Advisor Development program, preparing them to assist Senior Management in QI. The data capabilities of MaiKhanda team were also developed [XVII].

In addition to the QI trainings, custom-designed week-long management and leadership development workshop was also supported by THF in 2010 [IX].

6.5.6 Self Efficacy within MaiKhanda:

In the pre-intervention period, MaiKhanda was not able to engage district based stakeholders in QI activities especially the DHMT or District assemblies. The strategy was to work closely with both health centres and DHMT so that critical needs could be incorporated and budgeted for in the district Implementation Plan [XIV].

With the establishment of district based QI team leadership at district level in the late intervention period, involvement of DHMT in QI activities was perceived by MaiKhanda to be very effective. They felt that QI teams and DHMT owned the programme activities and worked together to improve the quality of care in their facility. However, a lack of budget allocation for QI activities in their DIPs (District Implementation Plans) affected effective implementation of the programme activities [XX], reflecting a lack of embeddedness of QI activities at the district level.

MaiKhanda perception regarding uptake of QI intervention by the facilities was usually positive but with very little data to substantiate the claims. For instance, during the early intervention period, improvement projects were perceived as instilling some kind of discipline among QI teams such that they took it as their responsibility to ensure that women coming to them for care do not die of complications that could be avoided [XVII]. However, there was no evidence to support this claim. Routine monitoring data within MaiKhanda was lacking throughout the length of the project.

During the late intervention period, according to MaiKhanda QI team, concerted support to QI teams in facilities yielded good results. Frequent coaching visits to QI team in the facilities motivated them. Health Centre and CEmOC Collaboratives had changed the mindset of many teams on the provision of quality care for the better as they continuously learnt from each other and developed new change ideas for testing [XV]. All this reflected an

increased confidence within MaiKhanda team to deliver QI interventions but, there were challenges in actual implementation on the ground. Frontline staff continued to be challenged by the huge expectations fuelled by the extreme poverty in the areas MaiKhanda worked in and the “hand-out syndrome” that has become so prevalent in Malawian society over the last 10-15 years [XXIII].

6.6 Individual Characteristics

This refers to the dynamic interplay between individuals in the health facilities and their organizations context that influences individual or organizational behaviour change. This section covers staff knowledge and belief about the intervention, their self-efficacy, motivation, teamwork and stability and leadership within the system.

6.6.1 Knowledge and beliefs about the intervention:

At the launch of the intervention in 2006, quality improvement was still a relatively new concept for the health care system in Malawi. While there seemed to be a better understanding of what the intervention was supposed to achieve, the links between the various QI activities proposed was less clear. A number of facilities suggested that they enjoyed the QI work and believed that it facilitated the formation of a network of facilities that were able to share ideas, knowledge and also resources. In fact this perception that new knowledge could be gained from attending QI workshops was an important motivating factor for the participating QI members [III].

However, facilities perceived difficulties in motivating team members and their colleagues to participate in QI work and attributed this to a lack of material and human resources [XII]. But at the same time, QI teams also felt that they could improve the quality of care at their health facility, despite the many constraints of material and human resources. While this was the belief of the staff, it was not clear how they would overcome the constraints. .

Some of the facilities felt that they were being pushed too much by QI activities, and felt frustrated at being steered too much into different directions.

One facility reported that they were not able to keep up to date with everything, and spread the changes to all the staff. This was because the changes were happening too quickly, in addition to the fact that there were too many changes occurring as part of the QI programme [XII].

From the CEmOC survey, more than half (52%) of the respondents felt that the QI process fits into their current system of work but with some effort and 12% of the respondents said that it fits in with great difficulty. Shortage of staff was the most common reason cited for the QI process fitting into the current working systems with some effort. About 44% of the respondents also felt that while QI was worth the effort, it was a little bit demanding [VII].

6.6.2 Self-efficacy:

In the pre-intervention period, there was not the same level of awareness of neonatal deaths as there was for maternal deaths in the facilities, and this perhaps resulted in lesser confidence to conduct neonatal death reviews in the facilities [XII]

According to the CEmOC survey, all the survey respondents unanimously responded that they were aware of data improvement activities happening at their health facility and that they were all involved in it. They were confident that appropriate records of births and deaths were being maintained at their respective health facilities. There were lot of improvements to the process of data collection, compilation and analysis but little progress in terms of reporting or using the data for decision-making. The use of run charts was not very common with the QI teams. Only 54% of the QI team members were aware of the use of run charts for analysis and decision-making. Even when there were improvements in data, they did not necessarily translate into data driven decision making [VII].

While there was a greater emphasis on profound knowledge, subject matter knowledge was not addressed in the intervention. On average respondents from the provider knowledge survey conducted in 2009 [V] , only got 70% of the questions correct with significant difference between staff at CEmOCs who answered 86% of questions correctly and staff at BEmOCs who got only

64% correct. Even simple knowledge such as the correct frequency of monitoring patients in routine labour was equally lacking in BEmOCs and CEmOCs with an average of only 36% and 33% correct responses respectively(213).

Knowledge regarding emergency newborn care was low with an average of only 58% correct answers. The biggest gaps in knowledge was about management of a baby who did not breathe spontaneously (only 35% answered this correctly) and management of a baby with low Apgar scores at 1 minute. Infact, 23% of respondents gave potentially life-threatening responses, including 8% who said no action was required and 4% who said they did not know what to do. There was a limited understanding of simple life saving procedures such as newborn stimulation and resuscitation, skin to skin contact and early initiation of breast feeding.

There was a lack of correlation between how recently a respondent had studied a specific subject and the likelihood that the respondent would get the answer to a question on that subject correct. Also, there was no positive correlation between the number of years of professional training received and the number of correct responses. Neither was there a correlation between how confident a respondent indicated that they were in a specific subject and how likely they were to get the answers to questions on that subject correct.

Individuals' belief in their own capabilities to execute a course of action was a poor indicator and unrelated to their level of clinical knowledge. Only 42% of the staff felt that the care they provided was ideal. Only 50% of the staff said that they would recommend a friend/relative to deliver at their facility.

While staff generally had a negative perception regarding the care they provide, this was more so in the CEmOC facilities as compared to the health centres. For instance only 27% of respondents indicated that their facility was always clean enough and only 42% indicated that patients got enough information to look after themselves at home. Interestingly staff perceptions seemed better in BEmOCs (48% ideal on average) than in CEmOCs (28%

ideal on average). CHAM facilities were better off in coming up with new ideas at work, as compared to MoH facilities [V].

By the late intervention period, the teams had tremendously improved in the way of reporting which showed a maturity and understanding in what they were doing. The format of presentations, the language and the graphical presentation all displayed an advanced level of understanding of the quality improvement in all the health centres [XXVIII].

6.6.3 Motivation

Intrinsic factors such as achievement, recognition, the work itself, responsibility and a sense of purpose seemed to have a far more important role to play in motivating people. The motivation for individuals to become involved in QI ranged from being in a team, to learning from others to being able to identify problems and find solutions. A number of individuals cited their reason to become involved in QI was to reduce maternal and neonatal deaths [VI].

The Collaborative sessions were considered motivating because these meetings and learning opportunities allowed the team members to gain knowledge and experience which they could take with them to benefit their own facilities [XII].

During the baseline period, facilities received input from LSTM's technical advisor, which allowed facilities to gain knowledge, and technical expertise. Feedback from a number of facilities revealed that the presence of the LSTM technical advisor greatly enhanced their learning experience when conducting their MDRs and that this was one of the motivating factors for the facility staff to complete the MDR forms [III].

General feedback on the QI programme was mainly positive, with a number of team members reporting that they really enjoy the QI work. MaiKhanda QI programme allowed for regular interaction between the facilities, and a number of facilities used the QI work as an opportunity to meet with clinical experts from other tertiary care facilities. These meetings and learning

opportunities allowed the team members to gain knowledge, which could be used to benefit their own facilities. The QI programme facilitated the formation of a network of facilities that were able to share ideas, knowledge and also resources [VI].

In addition, MaiKhanda programme was seen as raising awareness of quality issues in their facilities, when previously providers had not given any thought to it. For example, a QI team member from one of the health facilities explained that they had never prioritized hygiene measures in the facility prior to MaiKhanda QI work.

Motivation varied across districts and also by age of respondents. Lilongwe was one of the worse off districts in relation to staff motivation and patient engagement. Given its larger size, administrative management of Lilongwe has always been a challenge. Despite its geographical proximity to the Ministry of Health, it had a low staff to population ratio. Limited availability of resources (including human resources) was often cited as the reasons for low staff morale. Another similar reason for the shortage of staff for QI activities was the high turnover among the staff, which hindered progress in QI work [VI].

CHAM facilities were usually better off than MoH facilities in terms of staff motivation. The reasons for this could be the financial incentives related to working at CHAM facilities. While salaries for CHAM facilities were provided by the Malawi government, the staff got an additional 10-15% top up on their salaries. Further career opportunities were perceived to be better within CHAM organizations than MoH [VII].

The overall picture with regard to the will and motivation of staff to work is that, on average, staff are interested in their work and often willing to work beyond their job responsibilities if required, but were sometimes lacking in motivation which may partly be due to the fact that they are almost universally completely dissatisfied with the salaries they received [VI].

Lack of support systems such as drugs and human resources, high workload, lack of career progression were identified as the key demotivating factors.

Towards the end of the early intervention period, it was reported that Health Centre and CEmOC Collaboratives had changed the mindset of many teams on the provision of quality care for the better as they continuously learnt from each other and developed new change ideas for testing [XV].

6.6.4 Teamwork and stability

The QI teams were an integral and central part of the QI intervention, especially since MaiKhanda role in the intervention was facilitatory rather than direct implementation. MaiKhanda was dependent on the QI teams for the delivery of the intervention at the facility level. A number of teams were composed of a wide range of facility staff representatives including the Medical Officer, the Chief Clinical Officer, the Matron and the driver on duty. This would ensure that all the perspectives of care within the facility were considered, including transportation issues[XII].

The baseline evaluation observed no changes in the QI teams or leadership, unlike during the implementation phase of the intervention when having a stable QI team proved to be a challenge. This might probably be because the survey was conducted very early on in the project and health care providers were perhaps still excited with the novelty of the improvement model.

By the time of the CEmOC survey in 2008, 130 members had once been members of QI teams from mid-2006 to mid-2008, with 88% still retaining their association with the QI team. Another 11% were no longer QI members. Although only 15 of the QI team members dropped out during the survey period, there was anecdotal evidence to say that they were the most dynamic within the team and exhibited some exemplar leadership qualities [VII]. A lack of teamwork was present and linked to the human resource crisis. Because human resource was reallocated whenever necessary, each time new staff came in they had to be oriented.

Stability of QI team members also did seem to have an effect on QI work being carried out in the health facilities. Using number of PDSAs run as an indicator for the QI work in the health facilities, facilities with a greater number of PDSAs being run have had no drop outs in the QI team [VII].

Stability of QI teams was a huge challenge throughout the implementation period. One of the tertiary CEmOC facilities, for instance, had a completely new QI team, as the interventions were coming to an end in 2010 [IX]. Turnover of staff made organizational learning and memory weak and also hindered progress in QI work [XIII].

6.6.5 Leadership

QI programme hoped to introduce innovations such as leadership buddies, local experts, super improvers and learning other facilities through exchange visits. Except for a learning programme of super-improvers, rest of the ideas did not take flight.

At one point in the pre-intervention period, 70% of the quality improvement team had stopped carrying out their projects due to lack of support from senior management and a perceived lack of team spirit [XVII].

Leadership was not considered to be inclusive as staffs were less likely to be asked by their supervisors for suggestions and comments on work related issues [VI]. MaiKhanda conducted a 2-day learning session for leaders from Lilongwe health areas to orient them on QI and strengthen relationships between them and different levels of management at DHO level [XV].

Clinical and senior leadership were lacking in most health facilities throughout the intervention period. In places where leadership was strong and supportive of MaiKhanda's interventions, visible improvements were made. For instance, Kasungu in 2010 had the lowest performance of the teams in the second quarter of 2010. This was attributed to the change of District Health Officer, weak QI leadership, QI projects not being executed without MaiKhanda supervision, disjointed agreements on the agenda between MaiKhanda and what DHO want to achieve and how. On the other hand Salima performance

had improved during the same period, (for eg: 60 days without any maternal deaths in their facilities). The reason cited for this success was the appointment of one of the Clinical Officers who was also a long time QI leader, as the safe motherhood leader for Salima district [XXVIII]. The senior team at MaiKhanda felt that Salima district, mainly due to its activated leadership, was poised for rapid spread and would act as a catalyst for the other district managers to follow suit [XXX].

6.7 Programme Implementation

6.7.1 MaiKhanda's role

The programme execution was by the frontline staff (QI teams) within the health system. MaiKhanda's role was that of a facilitator. MaiKhanda's QI officers would conduct field visits to CEmOCs and health centres to ensure implementation of the change packages and the testing of new ideas using PDSA cycles by the QI teams. This intended role of MaiKhanda was based on the diffusion of innovation model where improvement is supposed to happen through influencing organizational and behaviour change at the local level. This strategy meant that the intervention had to have a lean management structure within MaiKhanda. Indeed, at the baseline and pre-intervention period there was only one QI officer covering the 3 districts and 9 CEmOC facilities. However, not enough attention was given to organizational context and health systems capacity to absorb these relatively new concepts in quality improvement. Eventually, there was need for a closer and much intense engagement from MaiKhanda to help build capacity around the concept of QI at the health facility level. While this was still a challenge, MaiKhanda itself as an organisation was evolving. It was transitioning from a project to be established as a local NGO within Malawi. This was coupled with poor planning and budgeting for the QI aspect of the project as a result of poor understanding of the existing health system. This however, stabilized in the late intervention period and more district staff were recruited in 2009. But staff capacity to understand and implement the QI concept was limited. It was overwhelming for MaiKhanda staff to develop a clinical understanding of the

situation at the same time apply QI practice to it and being very stretched across the district. Capacity building initiatives such as QI course and other efforts were made to bring the MaiKhanda team to speed with the current trends in QI [XX].

As mentioned above, the actual interventions were implemented by QI teams from within the health facility and consisted of collaborative learning sessions and actions periods. The collaborative learning sessions provided a platform for peer group learning. The QI teams met together on a quarterly basis during the collaborative learning sessions to share the ideas they had implemented and learn from other QI teams. The QI teams received support from the MaiKhanda officers during action periods to ensure that the change packages and shared ideas were actually implemented by the QI teams to improve changes in current practice in their facility. Actual implementation of this strategy on the ground was difficult. Firstly there was the constant mobility and dynamics of the health system, including staff shortages and staff transfers [VI]. Secondly, MaiKhanda staff were delivering the interventions through existing structures in the facilities where the mood for individual gains (incentives, per diems) superseded that of social gains [XXIII]. This was because there were so many stakeholders incentivizing the same staff to do so many of their own activities. Thirdly, some of MaiKhanda improvement advisors were non-clinical staff and this might have had implications for how they were perceived by front-line staff on the ground, given the hierarchical nature that exists within the Malawian health system. And finally, the concept was as new to the QI teams as it was to MaiKhanda staff and there was a steep learning curve, which had to be maintained within the constantly changing environment within the health facilities.

6.7.2 Programme Execution

At the start of the intervention there were 3 major implementation activities. These were, as described in the Theory of Change, data improvement, quality improvement activities which included a combination of criterion based audit and model for improvement and support for death audits.

Data improvement was one of the first and a key activity that was undertaken as part of the MaiKhanda intervention. Data clerks from the facilities were part of the QI teams and were involved in separate data learning sessions and received support from MaiKhanda during the action period. One of the key areas of capacity building was the collection, presentation and interpretation of data [XIV].

In the pre-intervention period, all the facilities, with the exception of Mlale and Salima, reported encountering problems with their QI implementation. These problems ranged from poor infrastructure to lack of personnel [XII].

Most QI work, especially up to the early intervention period, focused on maternal deaths and only MDRs were taking place, thus fostering improvement specifically in maternal health. However, some of the facilities under MaiKhanda project had initiated QI work for neonatal health in their facilities. For example both Salima District Hospital and Bwaila Hospital implemented patient attendants in the labour room, while CHAM facilities-Nkhoma and Mtundu developed neonatal sepsis checklists to help with sepsis diagnosis in the neonates [IX].

Acknowledging the discrepancy in neonatal focus, MaiKhanda had organised a neonatal expert meeting that was conducted by content experts in neonatal and paediatric care along with representatives from the facility QI teams.

Clinical knowledge regarding newborn health across the health system remained very weak for both essential and emergency newborn care. The intervention did intend to build capacity for clinical improvement without removing clinicians from the front line of care. Thus the focus was on in-situ training. In pockets where in-situ trainings were implemented, such as in St. Gabriel's hospital, improvements were noted in the competency and skills of the nursing staff in labour ward to conduct neonatal resuscitation [XXIX]. However, the intervention was not implemented at scale and despite attempts at clinical skills building both by MaiKhanda and other stakeholders throughout the length of the project, sustained progress was not achieved.

In the early intervention period, the revised plan for quality improvement consisted of the following components: Strengthened technical support to MaiKhanda and QI Teams, quality improvement in health facilities, clinical skill building, data improvement and leadership development.

With the slow progress of the interventions and the many logistical challenges of running the intervention, a suggestion for a shift from QI work in 3 districts to focusing on a single district i.e Salima was proposed. The idea was to make focused interventions in one district to show impact. *"Salima was chosen as it is the district most passionate about QI methodology, has made the largest commitment to the MaiKhanda work and has the highest level of DHO engagement."* However, this proposal was shelved mainly due to reservations from the evaluation team about preserving the integrity of the randomized trial study design [XXII].

By the middle of 2009, Health centres QI teams were largely working on tracking and managing high-risk pregnant women. Enhanced coaching visits was reported as being useful since teams were now not only generating data, but also using it to drive change in their facilities by continually testing new ideas. For example, 80% of health centres captured and used family planning data alongside data on deliveries – demonstrating a growing grasp of the whole systems approach. Good progress was noted especially with Salima team where they were able to present progress using run charts [XXV]. In Lilongwe QI officers managed a total of 200 against 229 planned visits in the second quarter of 2009 [XXII].

In the last quarter of 2009, there was a drop in planned versus actual coaching visits and supervisions (to only 50% in two districts in January). The main reasons behind the drop were the severe national fuel shortages in November and December [XXVI].

Overall, there was an improvement in the visits of the FI Officers to facilities. In the late intervention period, almost 80% of the planned field visits were covered by MaiKhanda staff [XXV]. It is interesting to note that even when the

programme was functioning fully well and was at its peak performance, 100% target was not met. This was partly because of over-planning- for time needed for attending Collaborative Learning Sessions was not included in the plan and attention was constantly being diverted to other needs, especially follow-ups on missing data and partly because of external circumstances beyond the control of the team such as rising fuel prices and fuel shortages affecting programme budgeting and implementation.

The visits generally focused support towards QI Teams in the health centres for identification of high risk mothers and following them up to deliver at the CEmOC facility, identification of potential blood donors before delivery for every high risk mother, use of referral checklist for every mother referred in labour [XXVI]. The use of the referral checklist and identifying high risk pregnant women was established in all the health centres in Salima, during this period.

In the late intervention period, QI work was still focused on reducing maternal deaths in the CEmOC facilities, mainly through an intervention to reduce 50% of deaths due to postpartum haemorrhage (PPH) [XXIX].

Change packages related to newborn care were not yet being implemented to scale. Where they were being implemented, the focus for most of the facilities was on reduction of neonatal death due to asphyxia by enhancing skills in neonatal resuscitation, reducing deaths due to prematurity by putting babies in Kangaroo Mother Care. A few facilities were using the neonatal sepsis protocol to reduce NNDs due to sepsis. The choice of the intervention was based on Pareto Graphs created by each facility to show their highest cause of neonatal death [IX].

The interventions gained momentum only in the last 18 months of the intervention (June 2009-December 2010). The intervention tried to touch up on all the critical success factors such as frequent coaching visits, team functioning, leadership support, learning sessions and refresher meetings. However these interventions were fragmented and manifest as isolated

pockets of excellence, rather than part of a coherent strategy. There was some achievement in neonatal sepsis treatment protocol and KMC protocol at Nkhoma hospital and this was attributed mainly to the clinical leadership available there, while there was evidence of improvement in resuscitation and deaths due to asphyxia in Mlale. The intervention produced isolated pockets of success in different health facilities on different parts of the intervention package. It was therefore difficult to establish if changes were happening independent of the intervention or if indeed it was influenced by the intervention [IX].

The project was originally designed to be scaled up to all areas of the three districts within five years, including the control areas designated by the evaluation design. This scale-up phase was postponed because the partners felt the existing set of interventions had not been successfully implemented.

In summary, by the end of the intervention period in 2010, the focus was still largely on CEmOCs and mainly on maternal change package. Neonatal change package had not started until 2009. Health centre collaborative were functional but, focus was mainly on referrals and no measurable outcome were established to measure this progress.

6.7.3 Finance

Although, great efforts were made to maintain a constant flow of funds for the project, in reality there were challenges to financial management of the project in Malawi.

During the baseline period (2007-08), the global rise in fuel prices had a negative impact on transport logistics for project activities. Fuel prices had gone up, impacting on MaiKhanda budget [XVII].

After the strategic review in 2008, there were changes to the intervention design, which also had implications for management. With LATH exiting the programme, the financial management of the programme was taken over by IHI. A new accounting package had to be installed [XV]. This transition brought about some delay in funding at the country level, which in turn

affected the timing of the Learning Sessions. As a result, some of the collaborative learning sessions conducted in the beginning of 2009 were done too close to each other without adequate time in between for action periods, to catch up with lost time. Also some of the collaborative had to be combined- the closing of 2008 Collaboratives and beginning for 2009 Collaboratives, which was not ideal [XXII].

In the late intervention period, the flow of funds was regular and much better managed. For instance reports from Salima, mentioned a constant flow of funds to the district which, brought satisfaction to the MaiKhanda facilitation team as they managed to meet their programme targets. This point was also echoed by the other two districts. District Coordinators presented a consolidated report (for facility and community intervention and for M&E) for their districts [XXIV].

6.7.4 Stakeholder engagement:

The QI intervention developed by MaiKhanda was in line with the national 'Road Map' for reducing maternal and infant mortality, developed by the Ministry of Health. But, the concepts of QI were still relatively unknown to the Ministry and other key stakeholders at the beginning of the intervention [XII].

All the respondents in the CEmOC survey valued the visits by MaiKhanda staff to their health facilities, during the action period. Majority of the respondents thought that visits by MaiKhanda staff were adequate, while 52% of the respondents suggested that they should visit the facilities atleast twice a month [VII].

MaiKhanda was perceived to have an influential and complementary role to play in the intervention by helping identify the problem area, develop plans and implement them

"…sometime you are doing things blindly but the way MaiKhanda has helped is to look back on data, analyse data and try to learn from data and implement what can make the situation better". (Nursing staff, St.Gabriel Hosptial) [XXIX]

Towards the end of the baseline period, there were some changes in leadership on the part of the Ministry of Health. This however, did not have a negative impact on the project. The new minister was very supportive of MaiKhanda project. But the challenge was the non-active involvement of stakeholders [XIX].

In the early intervention period, MaiKhanda was involved in the development of District Implementation Plans (DIPs) and budgets for the year July 2009 to June 2010 for Lilongwe, Kasungu and Salima districts for the first time [XXII].

During the early and late intervention period, at national level, a number of developments have combined to slow down the advocacy and stakeholder engagement efforts. There were also changes at ministerial level with the appointment of a new Minister and Deputy Minister of Health as well as transfers in the Ministry of Health (including key positions such as the Principal Secretary of Health). This meant MaiKhanda had to start building afresh, relationships at senior leadership levels in the Ministry. This also had an impact on the project implementation timeline [IX].

Chapter 7 Analysis

Having provided an overview of the project in the previous Chapter, I go on to analyse the program theory and implementation theory for the QI intervention. Analysis of program theory begins with a comparison of MaiKhanda's programme strategies with established mechanisms of successful QI intervention from the Michigan Keystone Project. The choice of Michigan Keystone Project as a comparative study was limited by the literature available at the time of analysis. There was not any successful QI study published, which described or hypothesized its QI intervention mechanisms. The only other study, which described intervention mechanisms, was the study by Greenhalgh and colleagues on modernization of health care services in the inner borough of London(217). However, this study focused on evaluation of the processes for whole scale transformation of health care services, rather than a specific QI intervention and was therefore not considered suitable for comparison. The choice of studies for comparison was thus limited and this in turn narrowed the scope of analysis.

Justification of this method of comparing studies with different context and outcomes is elaborated in Section §8.2.3.

In the subsequent section (§7.2), I try to develop an understanding of why MaiKhanda's programme strategies did not translate into intervention mechanisms. In doing so, I use a complexity checklist suggested by Ray Pawson(202) to outline and analyse MaiKhanda's implementation theory.

7.1 Understanding MaiKhanda's Program Theory

One of the very successful QI interventions described in improvement literature has been the Michigan Keystone Project, which was able to reduce central venous catheter, blood stream infections by about 66% over an eighteen-month period, engaging 108 ICUs in the process(173, 218). Mary Dixon-Woods and colleagues identified key mechanisms that influenced the

successful Michigan Keystone project(67). They articulated the success of the QI projects on:

1. Generating isomorphic pressures to join the programme and conform to its requirements.
2. Creating a densely networked community with strong horizontal links that exerted normative pressure on the members
3. Reframing the problem as a social problem with a strategic solution.
4. Using several interventions that functioned in various ways to shape a culture of commitment to doing better in practice
5. Harnessing data on infection rates as a disciplinary force
6. Using 'hard edges'

My first step is to compare if the mechanisms identified in the Michigan study are also present in the MaiKhanda study. In reality, the key QI strategies adopted by MaiKhanda were similar to Michigan study but that however, did not trigger a mechanism for the QI interventions in Malawi. In trying to understand why the strategies did not materialize into mechanism for improvement, I review in detail, the programme strategies adopted in the Michigan study and compare that with the programme strategies in the MaiKhanda intervention, to check for similarities and differences.

7.1.1 Generating *isomorphic pressures* to join the programme and conform to its requirements:

Institutional isomorphism is where institutes try to mimic each other. In Michigan, a collaborative model for QI in healthcare (similar to IHI's breakthrough collaborative series) was used in programme design which eventually involved 103 ICUs across Michigan all of whom agreed voluntarily to commit to the initiative (that is >85% of the ICU beds in Michigan joining the programme). According to Mary Dixon-Woods, isomorphism can be of three types: normative, mimetic and coercive. The ICUs within the Keystone Project exhibited either normative or mimetic isomorphism.

It is important to know the reasons why a facility would like to engage with a QI intervention. The health centres in MaiKhanda study had no say in their

choice of joining QI interventions. Their decision to be part of the intervention was approved by the Ministry of Health after the process of randomization done by the evaluation team. While in the Michigan study, ICUs voluntarily agreed to be part of the QI Collaborative (as a result of peer pressure), in the MaiKhanda interventions, the choice of participating health centres was decided based on their random allocation to particular arm of the cluster randomized controlled trial study design. This was approved by NHSRC (National Health Sciences Research Council) but, the districts were oriented on the randomization and the allocation of the intervention they were to receive, only after the process of randomization was completed.

On the other hand, the CEmOC facilities (which were not part of the cRCT design) were recruited through a process of orientation and stakeholder engagement. All CEmOC facilities within the three districts were engaged in the Collaborative.

Thus institutional isomorphism might have existed to some extent in the case of CEmOCs but not perhaps in the case of collaborating health centres. Even when it did, it is most likely to be through subtle coercion rather than normative or mimetic isomorphic pressures i.e coercive isomorphism-rules and regulations imposed by external agents. For instance, the decision to be part of the QI intervention was discussed first with the Ministry of Health and the district teams informed after the randomization process was complete. This lead to a perception as the intervention being externally developed and thereby influenced on how the model for improvement was taken up by the CEmOC facilities and the health centres. There were reports of health facilities feeling frustrated at being steered too much in different directions (§6.6.1). QI work was considered additional to the existing scope of work by members of staff and therefore lower in their priority list, affecting the ownership and therefore embeddedness of the intervention within the health systems in the long run.

The reason for isomorphic pressure not to have worked could also be because of the way the health system was targeted by MaiKhanda. Whereas

in Michigan study there was a targeted approach towards 103 ICU facilities, similar in form and function, in MaiKhanda, the intervention was rolled out across a vertical section of the health system. The health facilities organised across this cross section such as the CEmOCs, the district hospitals, CHAM facilities and health centres, all had very different capabilities. The structural and human resource arrangements across the facilities were different and a one size fits all approach would not have been the most ideal. For e.g. at the health centre level, a QI team would be a medical assistant and a nursing staff (this in many cases would also constitute the entire staff of a health centre). CEmOC facilities on the other hand had a multi-disciplinary QI team including data clerks and safe motherhood coordinator, since these facilities

Unit requirement for participation (reproduced from Dixon-Woods et al):

- Provide a commitment letter from the hospital CEO to the programme team
- Identify a project team leader, typically a nurse manager, who can devote approximately 10 percent time to this effort
- Form a multi-disciplinary project team
- Submit baseline and monthly infection rates data
- Complete a culture survey at the outset
- Participate in weekly immersion calls
- Participate in one or two conference calls a month
- Participate in a state-wide face to face meeting every six months
- Implement the programme's improvement tools

Box 4: Checklist for participating facilities in the Michigan Keystone Project

were larger and employed a range of staff. A replication of the 'Michigan model' would perhaps be best suited for the CEmOC facilities rather than the health centres. Furthermore, there were differences between district capacities as well. Salima, for instance, was considered an exemplar district

where there was perceived commitment from the district leadership and better women friendly care as compared to Lilongwe which was worse off in terms of staff motivation and workload (§6.6.3).

In Michigan study, participating hospitals had to make number of explicit but non-binding commitments signed by a hospital executive, along with a list of hospital team members and the amount of time they would devote to the project- See Box 4

In contrast, in MaiKhanda, commitments were made verbally and not signed by senior hospital executives. Infact, the senior hospital executives in the Malawian health system comprised of the District Health Management Teams (DHMTs) and they were passive rather than active participants in the intervention. As per the original intervention plan, DHMT were to be involved in planning and coordination of programme implementation and their leadership skills built (§5.2.2) but MaiKhanda was not able to engage them in QI activities in the baseline or pre-intervention intervention period (§6.5.6). Working with the DHMT as a strategy gained momentum only in the early intervention period (§6.3.4, §6.5.1, §6.7.4). While, involvement of DHMT in QI activities was perceived by MaiKhanda to be very effective (§6.5.6), there was a lack of budget allocation for QI activities in the District DIPs (District Implementation Plans) reflecting a lack of embeddedness of QI activities at the district level (§6.5.6)

It was also difficult to get a QI team to commit fully to the intervention given the staff turnover in the facilities and limited staff availability. Since there was limited staff capacity and it was usually the brightest that got taken away, it affected the team functionality as also organizational memory and learning (§6.6.4).

MaiKhanda's mentorship and coaching was spread too thin across the health centres and with a greater focus on the CEmOC facilities. Although programme implementation did improve over time, with more QI officers being deputed to the districts by 2009, it was challenging for the MaiKhanda QI

officers to support all the capacity gaps within all the 32 health centres and 9 CEmOC facilities.

Despite MaiKhanda QI interventions fulfilling most of the elements of the checklist (Box 4), they were unable to build the kind of isomorphic pressure mentioned in the Keystone Project. Some elements such as identification of a project team leader, getting a commitment of their time, completing a culture survey, or having a signed commitment letter from the hospital senior management, made the underlying assumptions and expectation for the Project explicit rather than implicit. It is likely that some elements, for instance, having a project team leader (§2.5.4) had a greater degree of influence in building isomorphic pressure among facilities, as compared to other elements on the checklist. This is not explored further in the Michigan study by Mary Dixon-Woods and colleagues. These elements were missing from MaiKhanda's engagement of the health facilities.

Another point related to the organisation of care and its links to isomorphic pressure is the organisational incentive to be associated with. Mimetic isomorphism in Michigan played an important role because it would have perceived it to *"be unacceptable or damaging not to participate in a particular programme"*(67), thus building a critical mass over a short span of time for the rapid uptake of the intervention and further scale up and spread. This is important in the context of American healthcare system where hospital reputation is linked to profits(219) However, such pressures were absent within the Malawian health system which are publicly funded and with limited resources for healthcare. The interest in quality of care and its continuous improvement is driven by different motives. Health care providers in health care settings such as the United States need to demonstrate high-quality care to compete in an oversupplied health care market. In publicly funded systems such as Malawi, health care providers need to improve quality to make scarce resources stretch further (17). Thus the broader health systems objectives also influence the kind of pressures that individual organizations will succumb to. This in turn, will effect the way the project is delivered at the ground level.

In MaiKhanda programme there was a greater emphasis on improving newborn outcomes using PDSAs whereas a focus on optimizing the limited resources would perhaps have been a better acceptable strategy for the health care providers who are front-runners of the health care system and have to manage within the resource constrained settings. This could be achieved by initiating Criterion Based Audits which builds on local facility performance, based on their local situational analysis. This was indeed part of the programme in the pre-intervention period.

Given how the intervention was organized in terms of the evaluation design, it would not have been possible to reach a 'tipping point' where it would be possible to exert isomorphic pressure. In Michigan, the spread was more like a ripple effect with more facilities joining the Collaboration over time, based on the principles of innovation diffusion theory (181), whereas in MaiKhanda participating facilities were spread out across the 3 districts based on random allocation.

7.1.2 Creating a densely networked community with strong horizontal links that exerted normative pressure on the members

In Michigan ICUs were asked to collect baseline data, participate in education and meetings through weekly teleconferences followed by a 2 day residential workshop. Submission of monthly data, monthly teleconferences and face-to-face workshop at 6 monthly intervals was also part of the intervention. The important aspect of the regular teleconferences, meetings and communication in addition to supplying information was to promote a networked, community based approach to the problem. However, the form evolved over time and modifications to the collaborative sessions (deviation from the protocol-referred to as implementation (in)fidelity) included 'cocktail hour' and networking session, a project token (such as a T-shirt), encouraging teams to lead sessions and present success stories, openly discuss problems and ways to overcome them, encouraging teams to learn form each other. A horizontal relationship among hospitals was encouraged leading to a learning community where people often spoke to each other outside the formal

structure. Collaboratives especially of professional communities help change professional practices by *taking their 'directions for performance' from inside rather than outside their professional group"*(67). They were also more likely to sustain collaboration and activity in the long run. Thus a "sense of community" in Michigan was fostered by opportunities for interaction and communication.

MaiKhanda programme envisaged the use of 'Knowledge Agents', recruited locally from the existing cadre of front-line staff for information sharing and continued learning and improvement and adaptation. They were to gather, aggregate and analyse data and act as the connective tissue for the programme, driving knowledge exchange and interpersonal connections and ensuring continuous learning. In reality, this was done by the MaiKhanda staff (QI officers) at MaiKhanda i.e 'external agents' from outside the health system. This was done during their 'Action Period' field visits to the health facilities. Submission of monthly data was thus dependent on the MaiKhanda staff making field visits and this was never a smooth process since there were financial delays leading to postponement of field activities (§6.7.3) and sometimes other extraneous factors such as fuel shortages, which were beyond the control of the MaiKhanda team, which affected data collection on the ground (§6.4.1). Even when MaiKhanda was at its best performance in the late intervention period, MaiKhanda staff were able to cover only 80% of the planned field visits (§6.7.2)

Even where data was collected, its use for analysis and decision making was limited. Most of the respondents in the CEmOC survey were not familiar with run charts, which is one of the commonest and routinely used method for data analysis in the model for improvement (§6.6.2). While data clerks were engaged in QI in CEmOC facilities for data collection, for the health centres data collection on key indicators was actually done by the M&E team, who were not part of the local QI teams or of the intervention.

It was only towards the later part of the early intervention period that health centres were showing great improvements in the area of data capture and use

(§6.7.2). The use of data and data as an engagement process, possibly happened towards a period when evaluation was coming to an end.

Monthly teleconferences in MaiKhanda were introduced in the early intervention period. However, the monthly calls mainly involved the MaiKhanda improvement advisors and external Technical Leads and not necessarily the front-runners in the hospitals ie facility QI teams. The purpose of the calls was mainly to build capacity of the project staff and develop a close network within MaiKhanda as an organization. There was participation from QI team members but this was not regular or well organized. In contrast, the network in Michigan study was with the individual QI teams from the participating hospitals.

Another important characteristic of the Michigan project in strengthening the network was the feedback mechanism- teleconference notes were recorded and available on CDs and a toolkit of materials to support implementation was distributed. In MaiKhanda, team engagement through regular teleconference calls was not practical since it would have been very expensive and was also logistically challenging to organize these with all the QI teams across the 3 districts. MaiKhanda did not have the infrastructure to support this. A greater emphasis at MaiKhanda was the on-site coaching provided to QI teams when MaiKhanda QI officers visited the facilities during the action period. This was greatly supported by the QI teams who felt that these visits were helpful especially in terms of supporting the teams with data from decision-making. MaiKhanda QI officers constantly reiterated to the QI teams, the purpose of the intervention, reviewing their data and discussing progress toward project goals. However, the coverage of these visits in the initial period was very low with one QI officer covering all 3 districts. Later in the programme in 2009, there were QI officer based in districts that had more frequent interaction with the district facility teams. Although on-site coaching visits gained momentum in the intensified late phase of the intervention, even then it was limited by contextual factors such as delayed fund availability and fuel shortages at national level.

The Collaborative learning sessions were held quarterly in MaiKhanda, unlike in the Michigan project where it was held bi-annually but, these were lead mainly by the facilitating team at MaiKhanda rather than senior programme leaders from within the health facilities as was the case in Keystone Project. Collaborative workshops apparently allowed for regular interaction between the facilities, and a number of facilities use the QI workshop as an opportunity to meet with clinical staff from other facilities [12]. Health facility staffs were interested and motivated from the learning sessions, as it “provided an opportunity to network and improve their knowledge”. Collaboratives seemed to change the mindset of many teams on the provision of quality care for the better as they continuously learnt from each other and develop new change ideas for testing (§6.6.3). However, there was little involvement of district based stakeholders in QI activities. This was a drawback from an improvement perspective, as some of the improvement activities required the engagement of district teams from the MoH. The Collaborative learning sessions had QI teams presenting at the storyboard session, but was not sustained throughout the length of the project.

The Michigan programme had features of a grassroots or "bottom up" movement. This approach was supported with a strong internal direction and 'top down ' leadership that empowered all stakeholders to participate. QI communities combined “*horizontal momentum with vertical integrating structure”* that not only coordinated activities but also potentially competing interests and motives. An important point of distinction here is that in the Michigan project, these vertical core structures were still part of the internal structure within the health system, unlike in Malawi where MaiKhanda was external to the health system and attempting to work through it.

In Michigan, they deliberately brought different stakeholder groups together. They targeted 3 groups: senior leaders, middle managers and staff of ICU. Workshops and teleconferences & other communications ensured that all voices were heard. In the Malawi programme, getting an inter-disciplinary team of people to come together and look at the problem collectively was one of the hallmarks of the project. This was successful in the CEmOC facilities

where there was a larger group of participants and less of a success in the health centres where the human resources was very limited. Again, this process was not led by the team leaders on the ground. In MaiKhanda, attempts were made to bring together different target groups. There was some initial success in this but, it was difficult to get hold of senior leadership to engage with the programme as, they were usually very busy (with various competing programmes vying for their time). Also the context in which the health system operates is different. Senior leaders for instance are not located within the same geographical space as the health facility staff. Distances exist and communication is very structured and hierarchical. Hierarchy in the workplace is very strong in Malawi. Staffs are generally reluctant to express what they think best fits their context. (§6.4.2.6). Thus participation in team meetings and collaborative was considered more of MaiKhanda's necessity rather than the local team taking ownership of these activities.

Finding the right team of project leaders to act on the "vertical core" was one of the highlights of the Michigan project. Unlike the MaiKhanda team, the Keystone project team seemed to have the right balance of scientific expertise as well as "*ability to engage emotionally with participants*"(67). The scientific expertise of MaiKhanda team was being built mainly through the external technical support and training courses (§6.5.5). The MaiKhanda team was on a steep learning curve even as it was providing technical support to the QI teams in the facilities.

While the Michigan team used a combination of hard and soft tactics to ensure discipline but at the same time gave participants responsibility and sense of ownership, MaiKhanda team had a facilitatory role to play and being external to the health system, it did not have the clout to influence or change clinical practices through an "insider" effect. The numbers of staff involved in the MaiKhanda programme were from a diverse background and with a stronger focus on QI. This could also have been a challenge for the QI teams in facilities to accept 'non-peer' groups as leaders or advisors to the interventions being considered.

Furthermore, MaiKhanda was not the only organisation, which was implementing interventions in these facilities, There were other organisations who were working on other aspects of MNCH as well (§6.4.1). In addition, there was also conflicting interventions such as the quality assurance programme run by other organizations that was closely linked with the quality technical working group at the Ministry of Health and thus able to exert a greater influence on improvement work happening at the district and sub-district levels (§6.3.3).

Despite all its activities, MaiKhanda failed to have a similar effect in promoting a networked community approach to the problem and developing a 'sense of community', as evidenced in the Michigan study. The key activities such as the monthly data collection, the teleconferences, the learning workshops and the coaching, all of which improved interaction and communication and thereby develop a sense of community in the Michigan gained implementation momentum only in the late intervention period in MaiKhanda. Moreover, the concept of QI was new to Malawi and the QI teams found the whole approach demanding of their time and effort (§6.6.1)

Some further exploration of the role of the human agency in the Michigan study could have helped to understand the intervention mechanism better. For instance, how many people from the Michigan project team was involved in reviewing the data and giving feedback through monthly teleconferences? Who were the people from Michigan leading on the calls and their standing in the medical community within Michigan? What was the power relationship between the call initiators and the call participants? Was it the same group of people participating in the residential workshops (exploring continuity of intervention)?

7.1.3 Reframing the problem as a social problem with a solution:

In Michigan, the intervention aimed at getting a community consensus on the problem as a social problem and this required: disrupting norms that treated central venous catheter-blood stream infections (CVC-BSI) as inevitable, developing a set of standardized interventions that the ICU community would

accept and implement. The programme was able to create something like "*a professional movement".*

While there has been significant gains in child mortality(10), the rate of mortality decline in newborn mortality in Malawi has been relatively slower(12). Maternal health has had more visibility over newborn health in the country. Malawi's problem of high maternal mortality was politically acknowledged, and the political leadership had the issue of safe motherhood high up on their agenda. For eg: Likuni hospital had a campaign for 100 free maternal death days. They organised a press conference where the Minister for health and other key stakeholders were invited, including the press. The high profile gathering helped project the problem as a social problem that could addressed through appropriate change packages being implemented in the facilities.

The QI strategy for newborns in MaiKhanda was to persuade health facility staff of the existence of a shared problem such as newborn or perinatal death around which they could organize. This was supported by developing and implementing a set of change packages around birth asphyxia, low birth weight and neonatal sepsis that the QI teams would accept and implement. The baseline evaluation recorded newborn change package had the most number of ideas to be tested along with data improvement, indicating a shift in the norms considering newborn deaths as unacceptable. However, in the pre-intervention and early intervention period, much of the focus was on maternal health.

Historically, newborn has been placed within the ambit of the Reproductive Health Unit whereas the sick newborn is covered by the ARI Unit within IMCI (Integrated Management of Childhood Illness) unit at the MoH. RHU does not have any newborn specialist clinicians or nursing faculty. Newborn is also under-represented in the emergency signal functions(188). The lack of health system capacity around newborn care is evident from the provider knowledge survey where only 58% of the survey respondents were able to answer correctly about emergency newborn care (§6.6.2). This usually prompts a

vicious cycle where clinical staff did not have the confidence to carry out certain procedures such as neonatal resuscitation, which eventually results in failure of the procedure and death of newborns. In-situ trainings were implemented in some facilities such as in St. Gabriel's hospital and improvements were noted in the competency and skills of the nursing staff in labour ward to conduct neonatal resuscitation (§6.7.2). Poor availability of resources was often cited as one of the reasons for not conducting perinatal death audits (§6.4.2.4). Although it was within these very constrained resource settings that maternal death reviews were being implemented.

Reframing a social problem is often met with resistance and Michigan used two strategies to counter this-story telling and hard data to overcome resistance and secondly, transform into a social problem the perception of CVC-BSI as a non-normal occurrence.

The MaiKhanda QI Collaboratives had story board sessions where facilities would present their stories using data collected by their QI teams. These were supposed to overcome any internal resistance and transform perceptions regarding newborn deaths as routine occurrence to an aberration. The capacity of facilities and QI teams within Malawi, to use their data and story effectively to transform perception and behaviour, was limited. Data collection in the Malawian health system is routinely collected to feed into the HMIS (Health Management Information System). A culture of reviewing data at the facility level and using it for decision making did not exist. As mentioned in the CEmOC survey, the number of people involved in data analysis and using it for decision-making was very sparse (§6.6.2). Although, data improvement was one of the change packages within the intervention, data collection remained a challenge for reasons mentioned above. Also, there was a greater emphasis on outcome indicators such as neonatal case fatality rate than on the processes of care. Focus on more proximal indicators of the processes of care (for eg: resuscitation rate) are more appealing to QI teams as they are able to relate to these indicators, take action and hold members accountable and these indicators usually have a story to tell. They also reflect the changes

in the facility influencing the processes. The distal indicators such as neonatal case fatality rates are usually influenced by factors beyond the control of the QI teams and can be de-motivating, if the results are not moving in a positive direction. This performance data (non-positive results) can be independent of the clinical improvement activities conducted by the QI teams- as increased mortality rate, for instance, can sometimes even be a function of improved data collection systems. It is therefore questionable if the storyboard sessions held during the Collaborative sessions, motivated staff and generated the kind of empathy that would instigate people to take action.

The transition from evidence to action in the Michigan project was mainly through peers with whom members of the 'community' could identify and by a process of reflection and discussion to arrive at a consensus. This provided members with identity and cohesiveness.

In MaiKhanda, it was the external technical advisors who reviewed most of the evidence for the interventions used in the programme. The MaiKhanda QI officers presented this evidence at various forums such as during QI collaborative and field visits, which was well accepted by the team but more as passive rather than active participants. Nevertheless, the programme would focus on a well-identified solvable problem. However, the use of evidence for decision-making was limited amongst QI teams and the health system in general.

As mentioned previously, the programme team (from MaiKhanda) leading the evidence based improvement interventions were not 'insiders' with whom the QI teams could identify. In places where MaiKhanda did manage to engage local leadership such as in Nkhoma hospital, these team leaders were expatriate doctors working in those setting and it is not clear if they shared the same relationship with local staff as other peers.

Identifying change agents and champions from within the health system was an important strategy within the intervention design of MaiKhanda so as to lead social processes that would influence change in the organizational

culture within the health system. But senior leaders from within the system were too busy to be engaged at a scale required by the project since there were also other competing programme which perhaps placed similar demands on their time [XIV]. Secondly, champions who were identified at the micro level either had competing programmes, which demanded their time, or attrition of this group of people was higher as compared to other QI team members. Thus there was the stable upper cadre in the human resource chain (such as at the MoH, DHO level) who could provide the leadership but were too busy to engage fully with the programme, followed by a lower cadre in the human resource chain (staff at health facility level) which was highly unstable since they were very migratory to engage fully with the intervention. Bwaila hospital, for instance, had a completely new QI team, as the interventions were coming to an end in 2010 (§6.6.4) Staff turnover was a huge problem and although the staff turn over proportion was much smaller (11%) during the intervention period, it was usually the high performing people that got moved, leaving a gap in the QI facility work (§6.4.2.5).

Similar to Michigan, MaiKhanda provided platform for critical reflection and discussion, mainly through collaborative learning sessions and action periods, which would enhance consensus and legitimacy. However, the major difference here was that in Michigan this was done 'within' the system with senior leaders from within the participating ICUs taking lead while in Malawi this was through an external NGO-MaiKhanda with a facilitatory role and limited decision-making authority. Thus, although the processes adopted in the Keystone project and at MaiKhanda were the same, in terms of bringing large number of people from many different organizations together and building a community consensus -the mechanism that worked in Michigan might not have had the same effect for the programme in Malawi.

Michigan Keystone Project had adopted a dynamic standardization process where key elements of the intervention were presented as 'essential ingredients' with a scope for local variation. This was a deviation from the original protocol which advocated the use of PDSA cycles. In contrast, in

Malawi there was a great emphasis on the methods i.e developing and reporting PDSA cycles. While local teams were innovative, they were required to report within the PDSA format which most often they did not since they found it to be time consuming. There was a greater push from the evaluation team to monitor the 'process' using PDSA tools. However, in doing so, evaluation failed to record the innovations that local teams were doing that could have been adapted by other teams during the collaborative sessions. Local adaptations did happen within the facilities such as the provision of cycle parking space for men accompanying the women to the facilities as part of the women friendly care change package. But this was not captured well either by the implementation team or the evaluation team.

Standardization has its own place in the organizational culture of the Malawian health system. Protocols and hierarchy are rigidly followed in Malawian health system. For instance, the Safe Motherhood Initiative from the MoH was not very appreciative or supportive of the changes to the labour graph proposed by MaiKhanda in its participating facilities, since it was not in line with WHO guidelines (§6.3.4). MaiKhanda also invariably ended up standardizing its interventions, mainly because of the difficulties of logistics management of such a large and diverse group of teams and secondly because of pressure from the evaluation team to maintain intervention fidelity.

Nevertheless, while MaiKhanda managed to put newborn health high on the priority list of health facilities in the baseline and pre-intervention period, it was unable to co-ordinate a kind of ‘professional movement’ in the early intervention period, similar to the maternal campaign it had rolled out in Likuni hospital. This could be linked to the organizational transition within MaiKhanda and the paucity of QI facilitators to build the momentum for newborn health, and the prioritization of newborn care within the health sector. With improvements in implementation ‘dosage’, there was a gradual increase in focus on newborn change packages in the late intervention period,

7.1.4 Change practice and culture at the hard end through interventions that function in different ways:

This meant achieving defined goals but at the same time also serving social functions. For example in the Keystone project, clinical QI interventions such as preparation of the surgical trolley (to reduce time) or use of a (locally adapted) checklist also fulfilled a social function. A fully prepared trolley served as an expression of the financial, logistical and operational capacity of the unit while checklists made visible the discrepancies between ideal & actual practice. *Checklist increased the visibility of individual contributions to the process … and made the process of CVC insertion into a ritual* (67)

Checklists were also used in the MaiKhanda programme. This was the referral checklist for mothers and newborns being transferred from the health centres to the CEmOC facility. But implementation remained poor. Data on implementation is also limited because of poor documentation. Criterion based audits, which were part of the baseline period but later discontinued in the project, also served the function of distinguishing between ideal and actual practice. Another intervention activity, which also served a social function, was the neonatal in-situ training and mentorship programme (§6.7.2). The neonatal resuscitation drill conducted at St. Gabriel Hospital required staff from the QI facilities to perform drill every week amongst all team members, making this into a ritual among the facility teams. In addition to improving the basic resuscitation skills and staff morale, the intervention also served a social function by bringing together all relevant staff at the same level but whether it had an effect in breaking the hierarchy, is something that would need further investigation.. Patients attendants along with nurses, midwives and clinical officers all took part in the resuscitation drill. Thus the checklist and the resuscitation drills could have influenced the restructuring of organisational and professional roles, relationships and identities within St. Gabriel. However, the intervention was not implemented at scale and despite attempts at clinical skills building both by MaiKhanda and other stakeholders throughout the length of the project, sustained progress was not achieved.

7.1.5 Harnessing data as a disciplinary force.

Systematic collection of data and feedback was a key feature of the Michigan programme. Data collection and analysis was centralized and feedback was provided to the individual teams. Data was visualized with posters and displayed in the wards. Feedback to the teams identified the data gap between programme goals and their current status. The team's also received anonymous report of their performance as compared to the other facilities in the Collaborative.

Data improvement and use for decision making is an important part of QI interventions. There was a strong focus in Malawi to use data for decision making and data improvement was one of the four change packages included in the model for improvement. Improvement in data required a two-part strategy that improved the quality of data *production* (the "push" strategy) and also improved the analytic *skills of potential users* of data for decision-making (the "pull" strategy). The data improvement change package at MaiKhanda tried to work around both these strategies. In MaiKhanda, primary data collection was usually done by the ward clerks and statistical assistants in the CEmOC facilities, while MaiKhanda M&E staff collected this data for the health centres. Separate learning sessions and on-site coaching and mentoring was provided to the data clerks by MaiKhanda staff. Data clerks were usually part of the QI team in the facilities. However, those collecting, processing and feeding back data to the participants did not have the necessary skills and resources for data management or analysis.

Most of the data analysis was done centrally and feedback on newborn case fatality rates were given to visiting QI team members during the collaborative sessions (as in the Michigan programme) as well as during supervisory mentorship visits to the facilities. But, this was not consistent during the project period with the national fuel crisis and internal programme funding delays affecting implementation activities and hampering long term engagement of team members. Further frequent staff turnovers, meant that feedback and follow up actions were also affected. It was also difficult to

establish ownership of data at the facility level.

While measurement of the infection free rates in Michigan study is likely to have encouraged "*tight coupling between infection rate and cultural change"*(67), this was not achieved in the MaiKhanda set up, especially for newborn care. Here the choice of indicators is critical. In MaiKhanda the key facility level indicator was newborn case fatality rate. While infection free days-the indicator used in the Keystone Project was within control of the hospital staff and was a reflection of their commitment, the effect of choosing case fatality rate as a measure of improvement had slightly different implications for the Q teams since all deaths were not within the control of the facility staff (for eg: women are brought dead to the facility). Thus infection free days (as observed in Michigan) is a measure of the process which is more within the implementers control whereas case fatality rate is a measure of outcomes and less within the control of the implementing QI teams at MaiKhanda. A greater emphasis on process improvement data such as proportion of babies receiving essential newborn care (for eg: immediate breast-feeding) or proportion of low birth weight babies receiving Kangaroo Mother Care is likely to have been well-received by facility QI teams.

7.1.6 Making skilful use of hard edges.

Michigan programme worked primarily through consensus and although it had no legal or formal authority, it did have some coercive features. For instance at ward level, rituals such as checklists increased visibility and procedural accountability. This gave the checklist a *'hard edge'. A second hard edge was the use of activists tactics to ensure cooperation. For e.g. ICUs failing to return data were subject to a number of sanctions by programme leaders(67)*

Michigan team had a lot of influence among the participating teams and although not explicitly mentioned in the paper, the engagement of the senior leadership, especially those leading the programme, might have had a significant role to play in this. On a parallel, the MaiKhanda team could not have the same level of assertion within the facility teams. Not only did they not have the power to exert authority, in many instances, they were infact

undermined. The Quality Assurance programme was operational over a longer period of time in Malawi and given their reach and relation within the Ministry (they were part of the Ministry's technical working group on quality) it was difficult to assert the Model for Improvement as a parallel or alternative, as suggested by MaiKhanda. The QI leadership within MaiKhanda did not have a clinical background and for a profession so much governed by the rules of hierarchy, it could be difficult on their level of organisational readiness to accept advice and support from a non-medical person.

7.1.7 Conclusion

The above section provides a summary of the programme strategies adopted by MaiKhanda and as can be seen, despite having programme strategies similar to the successful Michigan study, MaiKhanda did not 'trigger' intervention mechanisms similar to that Project. Some of the key factors emerging are MaiKhanda's role as an external agent, challenges to its internal setting, the external context of resource availability, challenges to engaging the health system, the health system's readiness for QI and scale up of interventions. While most of the strategies were implemented within different health facilities, the intervention did not manage to scale up to the proportions that was anticipated at the initiation of the project (§5.4). As the strategies failed to 'trigger' the mechanisms, I tried to explore further if the intervention was doing things right ie analysing its implementation theory in greater detail. This is the next section in my analysis.

7.2 Understanding implementation theory

Evaluating what caused a programme not to work is prone to misinterpretation. This can be avoided by reviewing the programme theory and implementation theory(220, 221). The distinction between programme theory and implementation theory is that the former identifies the contextual variables and examines the causal mechanisms that produce change while the latter relates to the justification and rationale of the programme ie uncovering the programme objectives and assumptions(222). Evaluating programme implementation theory is an important precursor to understanding programme mechanism and should be an integral component of theory based evaluation approaches.

An important point to consider in evaluating implementation theory is that interventions are not implemented as a linear model where the intervention effect can be determined solely by the dosage and fidelity of the implementation. Rather it is introduced into a complex health system which is built up of a large number of mutually interacting sub-units and exhibit complex systems properties such as synergism (i.e. the whole is more than the sum of its individual parts), self-organization and emergence (i.e. the whole cannot be reduced to the sub-units)(223). I explain this with an example.

Trying to fully implement all aspects of the change package was difficult within the limited resource settings within MaiKhanda and dynamic changes within the Malawian health system. Nevertheless, the interventions were implemented, but in a fragmented manner and manifest itself as isolated pockets of excellence, rather than as part of a coherent strategy. Bwaila hospital in Lilongwe, for instance, recorded a reduction in NCFR in the pre-intervention period even though they had not conducted neonatal death review, which was one of the key intervention strategies being facilitated by MaiKhanda then. But the QI team had worked on the neonatal change packages and tested four ideas related to referrals (3 PDSAs run), Kangaroo mother care (2 PDSAs run) and one PDSA for ideas related to correct use of

drugs and management of premature babies. On the other hand, Mlale and Nkhoma hospitals, which did undertake neonatal death review, did not have had a reduction in their NCFR. These facilities, incidentally also had less number of PDSAs being run in their facilities. In contrast, Likuni hospital which undertook neonatal death review, did have a significant reduction in their NCFR. In addition, they had also run about 22 PDSAs for four ideas which they wanted to test. This raises questions as to whether it was the nature of the facilities and their organizational readiness or the intervention characteristics of neonatal death reviews and change packages or a combination of these and other factors that influenced the NCFR in the aforesaid facilities. In terms of the implementation theory, there is a need to understand not only the implementation 'dosage' and intensity of the various intervention components such as death reviews, newborn change packages, data improvement but also their interaction with the health system.

It is difficult to establish if changes where reported, were happening independent of the intervention or if indeed it was influenced by the intervention. Linked to this is the issue of intervention attribution i.e if the interventions were making a difference and how much of it could be attributed to intervention efforts. While results observed in the MaiKhanda intervention areas were not significantly different than in non-intervention areas, its not clear if this is due to lack of impact of MaiKhanda in its target areas or due to MaiKhanda having impact beyond its target areas (as was the assumption at the beginning of the programme regarding the rapidity of the intervention spread) (§5.4)

Complexity is defined as "*systems display (of) behavioural phenomena that are completely inexplicable by any conventional analysis of the systems' constituent parts"*(164). In simple terms, the more difficult it is to describe the active ingredients of an intervention and their relation to each other, the more complex an intervention is. Complexity also needs to be understood as complexity of the intervention versus complexity of the system(224) although there is considerable overlap between the two.

I explain the implementation theory taking this aspect of complexity into consideration and use the acronym (VICTORE) suggested by Ray Pawson in his book 'The Science of Evaluation'(202). VICTORE stands for Volition, Implementation, Context, Time, Outcome, Rivalry and Emergence. I begin each section with a brief description of the acronym followed by an analysis in terms of MaiKhanda's implementation theory. I have re-arranged the variables according to their relative importance.

7.2.1 Time

Programmes change over time and intervention components that are shown to have an effect when a project is introduced, because there is impetus and novelty of the intervention, might not have the same effect towards the end of the project. For example while there was engagement of the QI teams in the earlier Collaboratives, absenteeism from QI meetings in the late intervention period prompted MaiKhanda staff to meet with DHO(Kasungu) and convince him of the importance of such meetings. [XXVIII].

The inter-relationships between the component parts of the intervention can vary the 'tipping point' of the intervention to either shorten or prolong the time-to-effect, sometimes rendering the intervention, premature for evaluation. For eg there were effects on late newborn mortality in the MaiKhanda interventions towards the end of the evaluation period and perhaps this was the beginning of the tipping point for MaiKhanda interventions. But the evaluation was completed by then and it was beyond scope of the evaluation team to measure these changes. Thus the time when a programme is being evaluated has significant effect on the interpretation of results.

A time-series study of breakthrough series quality improvement Collaboratives in low and middle income countries by Franco and Marquez found the average time to reach 80% compliance with standards or other measures of performance was 9.2 months and to reach 90% was 14.4 months(18). For traditional quality improvement projects Alemi and colleagues reported an average length of 17 months from identification of problem to completion of first pilot improvement while the average length of time to show effect for QI

Collaboratives, has been identified as ranging from 12 months to 24 months(225). The Michigan study (time series) showed improvement 18 months post intervention but, the outcomes were most significant in the latter half of the intervention period(218).

The breakthrough series collaborative meetings of MaiKhanda were held at 90-day intervals, over a two year period whereas in Michigan it was held biannually over a 18 month period. The rationale and evidence for this spacing remains unclear. It was perhaps influenced by the project time frame, but whether such short cycles are able to effect performance and outputs remains debatable. Despite the Collaboratives being implemented similar to the Michigan project, unlike Michigan there were several factors which influenced MaiKhanda's time-to-effect.

Firstly the pre-intervention phase of the programme was focused in the CEmOC hospitals. While the interventions in the health centres were initiated from 2007 onwards, it is not until the beginning of 2009 that the intervention activities were being conducted with regularity when there was a significant increase in "dosage" of support as a result of expanding the MaiKhanda team from one to six. Nevertheless, MaiKhanda time period was similar to other studies, for intervention effect.

Secondly, capacity related to QI was limited both within MaiKhanda and the field implementation staff at MoH. The key intervention components -the learning Collaboratives and QI team meetings in facilities was being implemented in parallel to capacity building of MaiKhanda on QI. There was no lag time between training and orientation of MaiKhanda staff and facilitation in the health facilities.

Thirdly, MaiKhanda evolved form being a project into a NGO and the process took about 2 years to have a larger local presence with increased ownership by the team on the ground. It was only around 2009 i.e almost 24-30 months after implementation of the Data Improvement package, when data started

getting reported into the reports [XXV]. It took this amount of time before data could be reported even within MaiKhanda.

From an intervention perspective, the setting up of MaiKhanda as an organization was a diversion from the core objective, and subsequently we see that the 'Model For Improvement' was actually implemented only from September 2008 to December 2010.

7.2.2 Implementation

The strength of an intervention implementation can be defined in terms of its dose and duration, intensity and specificity as well as implementation fidelity.

7.2.2.1 Implementation strength: Dosage of the intervention

The QI teams were regularly engaged by MaiKhanda, initially receiving a 'bolus' dose by way of "Breakthrough Collaborative Sessions", followed by a more 'sustained' dose during the action periods. Collaborative sessions for CEmOCs and health centres were held separately. They were visited by MaiKhanda field officers during the action periods, so as to facilitate the teams in thinking through the PDSA cycles. However, the field visits were hampered initially by non-availability of adequate MaiKhanda staff in early intervention period, especially at the district level and later by other external contextual factors such as the fuel shortages in Malawi and financial delays.

During the reflective evaluation (also referred to as the Data Deep Dive), the implementing team expressed their concern that improvements in one sphere of the intervention was counter balanced by failures in other parts of the intervention. This produced patchworks of excellence, which the programme was not able to sustain at scale [IX]. A system level interaction so that improvements in one area were not offset by failures in others was required. This was the dilemma facing MaiKhanda. There was a certain level of implementation 'dose' required to reach 'critical mass'. Knowing what that dose was, is important. Having a large scale of the project (usually without adequate resources) stretched existing facilitating (MaiKhanda) and implementing (QI) staff, thereby affecting the dosage of the intervention. For

instance one of the intervention districts Lilongwe, was bigger that the other 2 districts put together. MaiKhanda had little control over this decision, as this was decided at the level of the MoH [§5.2.3]. However, concurrent adjustments to the budget or implementation plan was not made to take this into consideration. On the other hand, having few health facilities with intense QI work may not yield desired results, since most of the interventions were linked to the 'vertical slice' of the health system and influencing improvements in the system requires a certain amount of scale. Finding the right balance of the optimum number of facilities and the optimum number of interventions was therefore crucial for the implementation of QI interventions.

The project lacked in capacity to follow through with some very innovative ideas such as the leadership buddy system, designed in the original proposal, where each facility would be connected to a higher level leader from the health system. The project was over-ambitious in that it had too many components of the intervention with limited staff from MaiKhanda to facilitate the activities.

Implementation was also influenced by MaiKhanda's capacity to implement the intervention. First, there was limited human resource available within MaiKhanda to implement QI interventions and when additional hiring was done later in the programme, it required time to bring them to speed with concepts of QI. Many efforts were made towards capacity building of MaiKhanda as a quality improvement organization, and staff were adequately trained in Improvement models. But the learning curve from being a learner to a facilitator was very steep, with improvement staff providing support to the facility team concurrently even as they were trying to absorb the QI concepts themselves.

Inadequate planning at the beginning of the project lead to limited human resources for QI being available within MaiKhanda ([XIV, XIX]. The design of QI intervention was not commensurate with the budget allocated to it. There was under-estimation of the resources (both human and material) required for programme implementation. This is a reflection of the technical partner's

assumption that key QI concepts would undergo rapid diffusion and be readily absorbed by the health facilities and teams without investment in human and material capital at the district level. This was realized and corrected by the early intervention period when QI staff were deployed to the districts.

7.2.2.2 Implementation strength: Intervention specificity

The specificity or conceptual clarity of an intervention is influenced by the length of its implementation chain. The longer the implementation chain the more it is prone to inconsistency and re-interpretation. The length of the implementation chain can be an important source of complexity.

The implementation chain in MaiKhanda involved a series of steps beginning with the engagement of the external technical advisors who would bring in their theory and advice and introduce it to the local team at MaiKhanda. This team had a limited understanding of QI concepts, only being introduced to QI as part of MaiKhanda intervention. They would then impart their skills and knowledge to the facility 'QI teams' through QI collaborative workshop and QI team meetings in the respective facilities. QI team members would then implement the intervention in their facilities where other non-QI staffs were also working.

One of the challenges was MaiKhanda's role definition. MaiKhanda was an independent organization facilitating the Malawian health system to improve their quality of service delivery and were not direct programme implementers. MaiKhanda's facilitatory role meant that they were largely dependent on the (already burdened) facility staff to take up QI on their own initiative. This was a challenge for successful programme implementation. While MaiKhanda's role in facilitating the intervention was appreciated, health care providers in the facilities felt that the QI models fits into their current system of work but with some effort. A majority of the respondents in the CEmOC survey felt that QI was a little bit demanding, but worth the effort [VII].

There were strategies to decrease the length of the implementation chain within the health system, by identifying QI team leaders and champions from

within the health system who would champion and lead the QI methods in their respective facilities. While this was a great initiative, challenges in retaining the staff persisted. Firstly in the larger hospitals there was transfers between the wards. Equally challenging were the staff transfer between facilities and district. As observed in the CEmOC survey, it was usually the most dynamic staff within the team that got transferred and they would usually also be the ones leading the QI work [§6.6.4, VII]. In places, where strong leadership was available, there was greater acceptance and adaptation of the QI work such as in Salima [§6.6.5, XXX].

7.2.2.3 Implementation Fidelity

One of the factors affecting outcomes was the implementation fidelity of the intervention. In public health interventions, it is difficult to predict this path given the involvement of the human agency at different stages along the pathway. This agency is not passive but makes its decision based on individual experiences and beliefs and is also greatly influenced by its organization culture. While this is difficult to measure, one simple descriptive method is to document intervention as planned and intervention as implemented. This documentation was lacking in MaiKhanda.

There were changes to the implementation strategy throughout the different phases of programme intervention. The key changes included shifting from an integrated model to an IHI prescribed model for improvement, shifting the strategic focus of MaiKhanda to be a QI organization from a MNH organization, decentralization of QI to the district level and greater engagement of the District Coordinators and towards the later intervention period, narrowing the implementation focus from 3 districts to a singular well-preforming district-Salima. These changes affected the implementation fidelity. Increasing the profile of the District Coordinators meant that there had to be greater communication and coordination of the central QI leadership with the district teams. communication remained a challenge between districts and with central office and as a result, stakeholders in the districts did not always see MaiKhanda as one project [6.5.4, XIX].

Should the implementation be considered of poor fidelity if the intervention deviates from original protocol? For complex interventions, it is likely that the intervention will deviate from the protocol and in most cases this is essential since the intervention is adjusting to the context such as the adoption of the in-situ training in MaiKhanda in the intervention period, based on the needs of the staff involved in QI, although this was not a component of the revised improvement model. But this poses a problem for measurement of implementation strength as it is not clear how the measurement scale would look like.

Adaptability (rather than conformity) is a key feature of QI interventions. However, even when there is agreement that intervention deviation from protocol is a good thing in the case of QI interventions, it still needs to be clarified, how far from the protocol can the deviation be and what degree of this deviation can still be considered as part of the program intervention. For instance, if the project had agreed to deviate from the original protocol and consider providing materials to hospitals or to go ahead and focus on Salima district rather than trying to cover 3 districts, would it still be considered as part of program intervention?

Furthermore, program fidelity is influenced by elements such as time. As time goes by the intervention matures and influences program fidelity. Even when intervention design remains unchanged and direction of causality is established, interventions can still vary over time and findings will be very sensitive to the point in time in which impact is measured. Impact over time can take a 'J' shaped curve. i.e. things get worse before they get better(35). For FI interventions, projects might have been on the J curve with changes in NCFR just beginning to show in the latter half of the intervention period. It depends on how long it takes for the QI concept to "sink into the system". Currently, there is no criterion denoting the appropriate time for evaluation of a complex intervention. Although studies are known to have shown results within 12-14 months and in the case of Michigan study it was 18 months (218, 225).

7.2.2.4 Implementation strength: Intensity of the intervention

The intensity of support provided by MaiKhanda's facilitation team was also an important factor affecting the implementation of the PDSA cycles in the health facilities. By the early intervention period, MaiKhanda was established as a local NGO within Malawi (rather than just a project, as was initially envisaged) and programme implementation through the facilitation team was encouraged with the Consortium members providing technical support where required. This transformation could have affected the intensity with which the intervention was being delivered on the ground.

Also the initial programme strategy was based on the innovation diffusion theory such that, a single MaiKhanda QI officer was facilitating the intervention across three districts, including Lilongwe which was a very large district. This influenced the intensity with which the intervention was delivered. While there is no documentation for earlier part of the intervention, it is clear from reports in the later intervention period that even when MaiKhanda was at full capacity ie having additional district based staff, it was able to complete only 80% of its planned activities [§6.7.2, XXV]. This was much lesser in the early intervention period (but documentary evidence on this is lacking).

There was an intense focus on micro-level application of the PDSA method without much efforts to improve the wider systems influence. MaiKhanda QI interventions by way of the Collaborative and Action Periods, was very structured and focused towards improving quality of health care in the 3 districts it was working. However, this was not supported by changes in resource availability to the facilities in these districts. This was a risk assumption that the project was aware of during its intervention design. There was indeed a change package within the QI interventions at MaiKhanda to consider efficient use of resources within its implementing facilities [§5.1.1.1]. But in a system where resources are scarce it is difficult to get facilities to prioritize, especially when MaiKhanda was external to the health system and did not intend to make a direct contribution to the resource gap. There was a change package dedicated to efficient use of resources. The change package

focused more on efficiency of existing resources rather than generating more resources. Attempts to bring in more resources mainly by the implementation partners was contemplated at one point, but the Programme Management Board (PMB) decided against it since that would not have been a sustainable model. Also there was a majority consensus that implementation fidelity needed to be preserved so as to compare the effect of QI intervention alone on quality of care in the absence of material resources.

The shortage of resources was linked to staff morale. As observed in the staff psychology survey, motivation was usually low and staff participation and engagement was also influenced by the hierarchal nature of the organization. There was little supervision (again limited by resources of human and logistics) and there was a culture of accepting directives rather than creative brain-storming [§6.4.2.6, §6.6.5]. Setting up of QI teams in the facilities was an attempt to challenge this status quo but maintaining and running QI teams till they reached a certain level of maturity to independently engage themselves in creative thinking, was itself a challenge given the staff shortages and the constant turn-over in the facilities. The QI teams needed a constant support from MaiKhanda's facilitation team throughout the period of the project. Also as seen from the CEMOC survey an environment fostering innovation and autonomy by staff to carry out improvement work was lacking, meaning that it would have been more difficult for them to test out new ideas using the PDSA cycles. The scope for creative thinking was limited since there was not enough support from the supervisor and other senior leaders in the hierarchy (§6.6.5). For instance, health facilities struggled to produce documents outlining their PDSAs.

The intense coaching that was to be provided throughout the period of the project, coupled with MaiKhanda's own internal transition, slowed down the progress of QI activities in most facilities that MaiKhanda was working in.

Other innovations such as developing local leaders had limitations because as was seen from the provider knowledge and skills survey, the knowledge &

practice related to MNH was very low among the staff despite the various trainings conducted by a range of stakeholders (§6.6.2). Their QI expertise was also limited, given that the exposure was very recent and mainly from IHI's learning sessions. Furthermore, as QI was a new concept within Malawi and therefore required an initial saturation of the concept (obtained through wide-spread orientation) for it to seep through into implementing teams. While much of MaiKhanda's staff efforts were focused on this, it was difficult to determine how or when this 'tipping point' would be reached.

There was little, if any, attempt to monitor MaiKhanda's work in the facilities and the intervention therefore was ill-equipped to analyse QI capability building amongst front line teams or the intervention intensity.

7.2.3 Context

Context usually comprises of the wider system level context as well as the local context. Here I refer to the context within the health system in Malawi and the broader political and social context. The context can be considered multi-layered consisting of different layers. Starting from the innermost layer and going outwards these can be described as individuals working in the health facilities, MaiKhanda relationships with the health facilities, organizational settings within the Malawian health care system as well as the wider policy and political context.

Beginning at the individual level, participants in the CEmOC survey were of the opinion that while QI was important to bring about improvements in MNH, it was time consuming and required some effort from their end. It is difficult to conclude from the available data whether the motivation to be involved in QI Collaboratives was influenced by the lure of personal incentives (such as per diems for attending workshops and meetings) or by individual's commitment to broader social gains (ie reduction in newborn case fatality rates in their facility) (§6.5.6, XXIII). Whatever the case, the capacity of the staff within the health facilities, in terms of their knowledge and skills related to newborn health was very limited. This is an important precursor for improvement since, it is the subject matter knowledge along with 'profound' knowledge that is

responsible for improvement according to Deming's theory of profound knowledge(46). While this was addressed explicitly in the pre-intervention period, through activities such as death reviews and criterion based audits, this was not given enough attention in the intervention period. In-situ trainings to address skills gap was introduced in the facilities and competency of nursing staff, for instance, increased after introducing neonatal in-situ drills at St. Gabriel's hospital. But the knowledge gap remained largely unaddressed in the intervention period and a clear strategy in this regard was lacking. There was also no clinical leadership within MaiKhanda with expertise in newborn care.

MaiKhanda's role as a facilitating agency has been explained above. This meant that the actual implementation of QI interventions was done by health facility staff working in the facilities, who were employees of the Ministry of Health with no formal affiliation to MaiKhanda. These staff were already stretched and their level of commitment towards QI might not have been as MaiKhanda would have expected. As MaiKhanda was external to the health system, peer influence would also have been very minimal. Furthermore, the facility staff was also participating and collaborating in various other MNH initiatives, which were simultaneously running in these districts. All these factors might have influenced their level of engagement and commitment to MaiKhanda's intervention.

Resource availability and staff turnovers were major contextual factors for MaiKhanda interventions. Essential supplies remained in short supply throughout the length of the intervention period as witnessed in the baseline survey as well as in the health facility resources survey . In a study conducted in 7 NHS hospitals in the UK, the systems reliability was about 81to 87% (226).This is much lower in Malawi (227). For instance availability of equipment in OT in a leading tertiary facility is around 63-67%. From the human resources survey at MaiKhanda, facilities had 1-2 signal functions available at any given point in time (33). In other words the reliability of the system was very low in terms of availability of resources. The degree of

organizational readiness for QI was also low characterized by the transient nature of the QI team members and the continued coaching that had to be provided throughout the length of the project so that QI team could grasp the concepts. Also, the hierarchal nature within the healthcare system characterized by a subordination and control mechanism that was prevalent across all levels including the staff-patient relationship was an additional challenge to promoting the model for improvement in these settings.

Every one of the nine CEmOC facilities in the CEmOC survey noted that a lack of material and human resources caused problems with QI implementation [VII]. Facility staff perceived difficulties in motivating colleagues to participate in QI work as a result of this. Shortage of staff led to an increased workload, which did not give the QI team members enough time to meet regularly and discuss QI activities. This in turn influenced their commitment to QI work as it was considered additional to their existing scope of work and therefore lower in their priority list.

Staff turnover was a huge problem because it took about 9 months for local sustenance of the intervention. Thus turnover of staff meant that there wasn't a consistent cohort to continue the interventions. For instance, a new QI team had only just been established at Bwaila only in 2010, one of the CEmOC facilities where MaiKhanda had been working since the beginning of the project in 2006.

"Because human resource is reallocated whenever necessary, so each time new staff come in we have to orient." "Since we are short staffed sometimes it becomes difficult to follow and do a protocol properly"[Clinical Officer, Bwaila Hospital VII].

Critical shortage of neonatal clinicians and neonatal care skills in facilities was also responsible for reducing attention to newborn care, e.g. Kasungu has not had a paediatrician for 10 years(§6.4.2.5, XIX)

Towards the early intervention period, there was a growing demand for skilled delivery against limited supply of quality services, expounded by external

factors such as the government ban on TBAs. With more women delivering at health facilities, the capacity of health facilities was severely stretched [IX]

In addition, other system level contextual factors such as the influx of 'global' HSAs also happened within the timeframe of MaiKhanda programme. The influence of these external factors on the attribution of MaiKhanda interventions to achieve improvements in newborn mortality is complex.

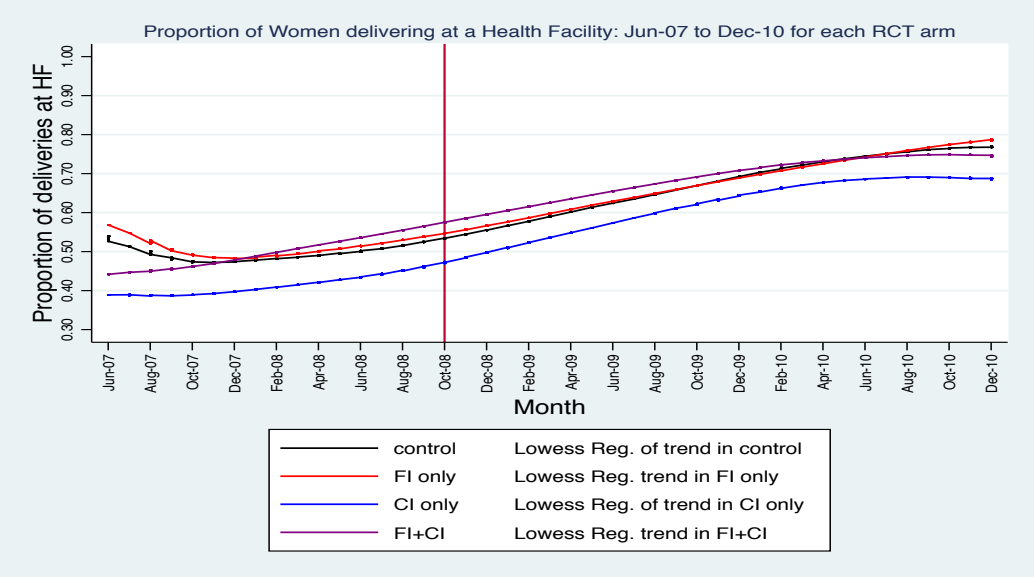

Figure 22: Increase in facility delivery in MaiKhanda (Source: Evaluation Report)

Firstly, there was an increase in deliveries across all health facilities (Fig) and this increase is more likely to be related to enforcement of the government ban on TBAs, by local leaders than any of the intervention change packages. Secondly, improvements in some aspects of the QI intervention could potentially have been masked by the sudden influx of patients coming to the facilities for delivery.

7.2.4 Volition

Social interventions do not work by themselves, such as in a drug trial, but rather it is the interpretation of the subjects implementing the programme, in this case the QI teams, that produce the results and these subjects are active agents not passive recipients. Their understanding of the interventions is

greatly influenced by how the programme theory is introduced to them and the interactions they have throughout the programme implementation with the different contextual elements in the health system.

MaiKhanda came into existence from a consortium of partners based in London and Boston. While the Consortium agreed on the larger role for local leadership, it was not until two years into the programme that due attention was given to this issue. Even when MaiKhanda got registered as a local NGO in 2008 and had a Director, the technical leads (represented by various consortium partners) were not necessarily engaged with the senior leadership at MaiKhanda, often by-passing them to work directly with QI programme staff at MaiKhanda. This gave rise to tensions within the organization. The root cause of this tension lay in the conceptual understanding of the programme. While MaiKhanda had a supply and demand side intervention, IHI considered this a quality improvement intervention and deemed itself to be the lead partner with principles of QI overarching all programme interventions in MaiKhanda. They were supported by the donors in this decision. This difference in conceptual understanding had implications for the intervention design and subsequent implementation.

The rising tension between the different partners manifested as follows. Partners were of the opinion that MaiKhanda had the mandate for decision making, yet they were not using it for decision making. There were also issues around flexibility and attitude of the Director and her capacity to adapt QI as well as issues around not being able to engage MaiKhanda's Board. At first instance these appear as management issues, yet the underpinnings of these lie in the programme focus and approach. The fundamental question that remained unanswered was if MaiKhanda was a QI organization or if it was a maternal and newborn organization This was never resolved neither was a shared common vision reached. As the programme evolved, the issues started manifesting as management or personality issues. This had implications such that, as MaiKhanda the community based women's group empowerment model and advocacy had relevance within MaiKhanda's scope

of work. However, as a QI organization, the focus of engaging women was merely to get them to the facilities where they would be managed appropriately (using QI methods).

Secondly, MaiKhanda was a facilitator rather than a direct implementer of the intervention and this also had implications of how the programme was articulated, among different levels of stakeholders within the programme (Figure 23).

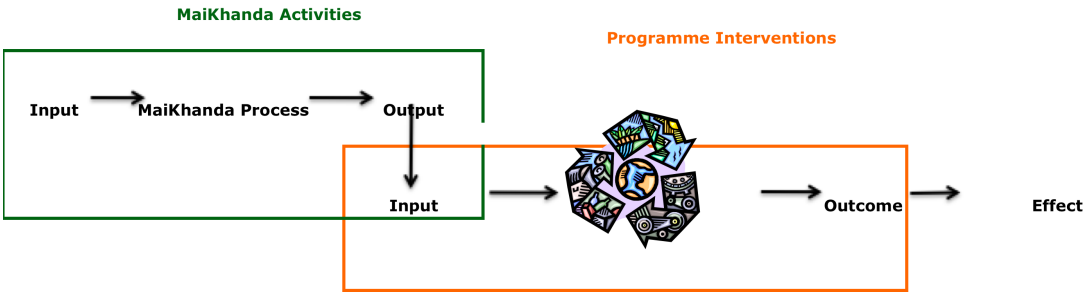

Figure 23: MaiKhanda's role as facilitator in the intervention

MaiKhanda's programme theories were generated by external change agents such as the Consortium partners which had first to be imparted to MaiKhanda staff who (the assumption is) would have accepted the theory of improvement. They would then introduce this concept to selected personnel from different cadres of Ministry of Health forming the QI teams, who would go on to implement the intervention at the health facility level. Initially the idea was to engage different stakeholders from various categories in the health systems through the leadership buddy system and knowledge agents. However, in programme implementation phase, it was only the QI teams (i.e. facility level staff) from the health facilities that the MaiKhanda programme closely engaged with. At the health centres level, these QI teams included the medical assistant and nursing staff who were often the only clinical staff available for facility, round the clock. Nevertheless, the assumption here was that the health facility staff would easily buy into the theory and implement it, within the context of a myriad of other similar interventions also vying for the same staff to implement their respective programmes.

The intervention theory did not take into consideration the existing perceptions related to quality improvement and quality of care amongst the staff in the facilities. The concept of QI was relatively new to the facility staff and introduction of improvement intervention was subject to their interpretation of quality of care in their setting. This influences their choice in engaging with the QI interventions. The collaborative workshops objective was to develop a collective understanding of quality improvement and how it could be implemented in the participant's respective settings. However, given the limited exposure health facility staff have had to 'the ideal settings of care provision', a good starting point could have been the MoH approved interventions such as Criterion based audits or death reviews and then use them as a channel to introduce the Model for Improvement. Although part of the initial intervention design, this was not part of the early or late implementation period.

Along with the processes of care, quality improvement at the microsystem (health facility) level also involves the content of care and these together are a crucial and integral part of continuous quality improvement (CQI). The focus of content of care is towards developing and advancing subject matter knowledge (the knowledge of medical sciences) and ensuring that standards and protocols are well understood so as to provide atleast the minimal level of care to all beneficiaries. In well-developed health systems, the underlying assumption is that there exists sufficient subject matter knowledge (content of care) within the health systems and that standards and protocols are well in place. Therefore in such settings the emphasis on quality improvement is focussed towards improving the processes of care. However in the context of Malawi, there needs to be an equal emphasis on traditional quality improvement methods as much as we would have on the processes of care. The assumption here regarding subject matter knowledge (i.e medical sciences knowledge) among health care providers is good and that standards and protocols are in place and being stringently followed, does not hold true. The lack of subject matter knowledge was evident from the provider knowledge survey. This was also the experience of the evaluators as well as

the TAs who had been conducting maternal death reviews in different health facilities in the different districts in Malawi. Content knowledge thus needed to be enhanced by repeated and continuous training on the essential elements of care along with training and mentorship for quality improvement.

7.2.5 Outcome

An intervention usually has multiple outcomes all of which would ideally need to be captured within the evaluation, but is seldom possible for any one evaluation to cover all the multitudinous outcomes.

The tendency of the impact evaluation of MaiKhanda, using conventional evaluation methods was to standardize interventions so as to maintain intervention fidelity. It did not establish feedback mechanism for programme implementers to make mid-course correction. However, dynamic change is an important characteristic of an improvement model and maintaining the fidelity of the intervention posed a major challenge to measuring outcome.
MaiKhanda intervention had a range of outcomes and effect indicators, the most important ones being captured in the process evaluation (Figure 6).
Linking the outcome to particular aspects of an intervention remained a challenge. It is not clear if the increase in health facility deliveries observed in the intervention period, was a consequence of the interventions such as women friendly care or an influence of the external context such as a ban on TBAs. Whatever the reason, the influx of mothers to deliver at the health facilities without addressing the issue of staff shortages lead to increase burden on the already over-stretched staff to provide skilled delivery at birth.

Choice of outcome in the initial stage is important. Working in QI areas where the outcomes are less sensitive to behaviour change, providing positive feedback to the teams is difficult. For instance, project was focused on measuring neonatal case fatality rates from the beginning using SPC (Statistical Process Control) measures. Statistical process control is adapted mainly from the manufacturing industry and is best used to measure changes in processes (72). Reporting and feedback on distal outcome (rather than process) indicators such as NCFR (Neonatal Case Fatality Rate) is likely to

have far less influence on the QI teams than the visibility of proximal process data such as use of partograph or skill score for neonatal resuscitation drills. Analysing outcome measures on a monthly basis can only offer little in terms of systems insight and is likely to cause more dissatisfaction and de-motivation for the QI teams. Rather, more tangible and proximal indicators such as the use of partograph could readily appeal to the QI teams. Along with visibility such local level process data also influences the micro-level point of care which QI teams have more control over rather than higher levels of the system.

Also from an evaluation perspective, it is difficult to associate changes in case fatality rates with improvements in service delivery as it is only one of the factors that influence case fatality rates. Non-linearity is an inherent property of any system and it is difficult to establish direction of an intervention(228). For instance, a study at St.Gabriel hospital, on improvements in neonatal resuscitation, showed a fall in asphyxia case fatality rate in late 2010. However, this was unrelated to a concomitant change in neonatal case fatality rate at St. Gabriel during the same period. Similarly, it is difficult to link PDSA cycles (QI intervention components) to improvements in service delivery. Context also has a bigger role to play. In addition, other system level factors might also influence the intervention.

From the MaiKhanda interventions we have seen that changes in CFR do not necessarily co-relate with the introduction of the neonatal change packages (Figure 17). However, as part of the improvement efforts for maternal health, QI concepts were introduced into the facilities which included general system wide improvement efforts, for eg: data improvement, which might have been taking place and could have had an indirect influence on newborn case fatality rate even before the neonatal change package was implemented.

According to Stames, evaluation implemented at local level does not immediately inform evaluation of the impact of the overall programme. This is because, while the overall programme might have a common causal theory, particular sites may have different implementation theories of how to activate

this causal theory(229). As reported by the Technical Lead from MaiKhanda [XXXI] during the intervention period, there have been islands of excellence within the MaiKhanda programme such as the resuscitation drills, use of partograph, referral systems. But the impact of this (defined by facility NCFR) on the overall programme could not be determined perhaps because the critical mass to impact on the common causal theory was not achieved. The programme would have benefitted if evaluation could dwell deeper into the implementation theories in these pockets of excellence, alongside various process implementation studies.

There are also unexpected outcomes. Although FSB (Fresh Still Birth) was not a primary outcome to measure, there were reductions in FSB rates in the health centres. This was an unexpected outcome of the intervention. Perhaps the maternal change packages that were more prominently implemented throughout the intervention period is more likely to have had a greater influence on stillbirth outcome as compared to neonatal case fatality. Although neonatal change packages were implemented, this was not until the late intervention period.

7.2.6 Rivalry

Programmes are usually implemented alongside existing interventions and therefore are competitors. The QA (Quality Assurance) programme by JPHIEGO was already in existence when the MaiKhanda quality improvement interventions were being rolled out. This produced a challenge, since they were part of the MoH quality technical review group and had different conceptual stand-point regarding ‘quality improvement’. The Ministry through the Quality Technical Working Group was more oriented towards a ‘Quality Assurance’ model and thus focused on standards and audits. Infact, MaiKhanda intervention did have standards and criterion-based audits as part of the programme component in the pre-intervention phase. However, the inclusion of the Model for Improvement in the intervention design for the project was controversial from the very start. The intervention design based on the IHI model for improvement included implementation of key change

packages for improving newborn mortality with the use of PDSAs by facility staff supported by an improvement advisor. The fundamental difference between the two models was that CBA was based on the principles of quality assurance and thus concerned with maintaining minimum standards, while the Model for Improvement was based on the principles of quality improvement, where focus is on constantly evolving and improving the standards of care.

The re-design of the intervention in 2008, ended up with the exit of LSTM the key implementing partner for standards and criterion based audits. On hindsight, this could have influenced the uptake of QI interventions more promptly by the facility based staff, as the line of thinking with the Criterion Base Audit was more closely aligned with the 'Quality Assurance' model that the Ministry was promoting.

It was not only internal rivalry that MaiKhanda had to deal with. From the survey of the MaiKhanda CEmOCs in 2008 it is clear that the MaiKhanda facilities intervention was not done in isolation and faced competition from other projects that were also in operation at the CEmOCs, vying for the same human resources. An average of two other projects were also concurrently in operation in each of the CEmOCs in 2008 [VII]. This lead not only to the competing demands on health care workers time but also of competing influences on the health outcomes and the resulting problem of attribution to specific interventions.

7.2.7 Emergence

Programme components come together to produce another novel component that actually changes the conditions that make them work. This component cannot then be broken down into its individual sub-units.

Criterion based audits and death reviews were part of the design in the pre-intervention phase. The death reviews were conducted by staff from LSTM and MaiKhanda. The expertise of the clinical mentors from LSTM during the death reviews gave rise to impromptu clinical skills mentorship during site

visits. Clinical mentorships became an important component of the pre-intervention period, later being replaced by in-situ clinical training sessions in the early and later intervention periods. These were emergent properties of the intervention that was not part of the original intervention design.

MaiKhanda started off as a project supported by external technical support to bring in the model for improvement within health facilities in the central region of Malawi. To gain legitimacy and acceptance and accelerate the QI work in Malawi, the PMB (Programme Management Board) supported the establishment of MaiKhanda as a locally registered Malawian NGO. Thus what started of as a project, evolved into a local organization, gradually building on to its own set of culture and value practices. But this was more demanding and time consuming than expected and not part of the original proposal. Much effort was directed towards getting MaiKhanda established as an organization and this affected the core functioning of the intervention which was to impact on newborn health in Malawian health facilities using Model For Improvement techniques.

There was a lack of senior leadership within the programme from the start. This was acknowledged and corrective steps were taken to get a Director on board. However, the skill set of the senior leadership was geared towards establishing MaiKhanda as a NGO and focusing more on advocacy for MNH. The technical advisors were more in alignment and working closely with the FI Team Leaders bypassing the senior leadership thereby leading to conflict between the local leadership on the ground and the technical lead on the project while the programme struggled to deliver on its core functions[§6.5.4, XXXII]. One of the ideas emerging from this was a proposal to focus work only in Salima district, since Salima district was perceived to be more engaged and involved in the QI work. The strategy was that quick gains there would trigger other districts to have a pull factor towards QI interventions. This was however never cleared by the Programme Management Board because of strong pressure from the Evaluation team to maintain the fidelity of the intervention for the purposes of the evaluation design.

7.3 Conclusion

In summary, most of the programme strategies associated with the mechanism of successful QI interventions, that were observed in the Michigan study were also present in MaiKhanda programme. Yet we do not see the translation of those programme strategies into 'mechanism' within the MaiKhanda programme. I then looked at actual programme implementation to explore how programme implementation might have influenced the mechanism of the intervention.

Similar to the Michigan project, MaiKhanda project also demonstrated different levels of success and contextual factors might have had a key role to play in it. For instance, availability of resources in Salima was relatively better as compared to Lilongwe and Kasungu. This had implications for staff motivation as well as uptake of QI interventions. The leadership was also reported to be very committed to improving maternal health. Thus the picture of QI interventions in Salima were different from those observed in Kasungu or Lilongwe

MaiKhanda was a project, initially known as the THF consortium, entailed with the task of implementing a supply and demand side intervention in 3 central districts of Malawi with the aim of reducing maternal and newborn mortality and morbidity. However, it later on registered as a NGO with its own governance structure. This transition did consume a lot of the time of the senior management team and diverted their attention from the core components towards governance and institution building.

Another major challenge was the internal capacity within MaiKhanda to facilitate the QI interventions. Staff were being trained and capacitated in QI as interventions were taking place simultaneously.

Death reviews and criterion based death audits which formed the basis for initiating improvement in the facilities, in the pre-intervention period, were not part of the intervention design in the re-defined model for improvement. As death reviews were part of the roadmap it could have larger buy in from the

staff, stakeholders on the ground to implement, which could have been the foundation stone for Model for Improvement work and also as a strategy to make in-roads into health facilities where there were other competing projects.

Given the wider system influences, it is difficult to link PDSA cycles or Collaboratives with outcomes

The project observed an increase in health facility deliveries. It is not clear if this was a result of contextual factors such as the government ban on TBAs or an attribute of women friendly care change package implemented as part of the Qi intervention or a combination of both.

The emphasis throughout the project was on the model for improvement and implementing it as described in the protocol rather than considering the MFI as a framework for implementing QI within the Malawian context. For e.g. the evaluation emphasized an adherence to monitoring the PDSA cycle. Given the varying context, it would have been prudent to focus on the theories of the intervention rather than the focusing on the key intervention activities per se.

The dosage of the interventions was probably adequate but the duration and intensity of the QI interventions were perhaps not optimum to trigger the intervention mechanism. There was also a lack of conceptual clarity from the very inception of the project. Long implementation chain of the intervention and intervention and the emergence of MaiKhanda as a local NGO also affected the programme implementation.

Chapter 8 Discussion

This PhD describes the evaluation of a quality improvement intervention to reduce newborn mortality in three central districts in Malawi. I begin this chapter with a summary of the key findings and then go on to discuss some of the emerging methodological issues related to evaluation of QI interventions. Following this, I outline the key learnings from the study that can contribute to future evaluation designs and highlight the scope for future evaluation designs involving QI interventions in resource poor settings. In final section, I outline the limitation of the study.

8.1 Key findings

8.1.1 Choice of improvement theories

MaiKhanda's programme theory for quality improvement is based on Deming's theory of profound knowledge, where subject matter knowledge is complemented with profound knowledge(46). The scale up and spread of this improvement is hypothesised to happen through a technique known as Breakthrough Collaborative Series. The theoretical underpinnings of this spread model are based on the diffusion of innovation theory where the rapid spread of new ideas or practices happens largely by imitation(182) mainly through the (human) agency categorized as innovators, early adopters, early & late majority and laggards(192) (§3.4). Based on the diffusion theory, the idea behind the MaiKhanda's collaborative workshops was to provide a platform where informal networks would be established alongside information sharing and the influence of opinion leaders and change agents harnessed. By identifying the innovators and early adopters of the intervention from the QI teams and promoting and building a conducive environment, it is assumed that scale up of these ideas will take place at a rapid pace. Innovation and scale-up happen almost simultaneously or parallel to each other. One of the reasons for this parallel approach is because of a whole systems engagement, that is, involvement of the human agency and other inter-

dependent contextual factors of the health system that facilitate the intervention outcomes (although they are not direct components of the intervention package)(95). Examples of newborn interventions led by IHI in the Tula Oblast region of Russia(230) and the Ghanaian Fives Alive project reflect such a model of spread (19). Sustainability and spread are thus an inherent property of the Model for Improvement.

Theories are both formal and informal. The use of formal theory in improvement such as Deming's theory of 'profound knowledge' (46)and Roger's 'diffusion of innovation' theory(181) guiding MaiKhanda interventions, help in learning and understanding of how improvement can work (or not) across a range of context.

For a program such as MaiKhanda that deals with multiple layers of the health system, theories of organizational behaviour change are crucial to understand health system dynamics but need to be considered alongside theories of individual behaviour change and social network theories of collective behaviour(183). Theories of organizational culture such as 'competing values framework'(108)are equally important to understand the culture and thereby the context in which collective behaviour is influenced. Other theories such as the Normalization Process theory, which help explain social processes that lead from an interventions inception to its normalization within a system(183), could have been considered over the diffusion of innovation theory.

One of the main challenges in theory formulation has been on deciding the rationale for the choice of theories as there is no set prescription for the selection of theories. A wide range of theories are applicable for an intervention ranging from agriculture, medicine, public health, organizational behaviour, psychology, political sciences, marketing to name a few(183). There are more than 60 theories, models and frameworks relevant to implementation available to practitioners & researchers, providing a bouquet of theories to choose from. There have been recent attempts in coming out with a taxonomy of theories for implementation research(231) and more specifically for improvement research(232).

8.1.2 Developing Theory of Change

The 'best fit' theories can be determined by developing a comprehensive programme theory based on the logic model of the intervention. Development of the logic model is considered an important step in theory-based evaluation. However, developing a logic model with all the underlying assumptions and rationale (also referred to as Theory of Change) is easier said than done and literature on how to do this is only slowly emerging now(233).

The challenge is partly because interventions are very context specific, implemented within a complex system and it is therefore difficult to prescribe a generic logic model, which can be applicable across a wide range of interventions. The other challenge is in establishing objectivity in the development of a logic model so that scientific rigor can be maintained. Nevertheless, the 'Model for Improvement' and 'Breakthrough Collaborative Series' provide a logic framework for MaiKhanda's interventions. These logic models (with slight variations) have been used in other settings, such as the Keystone project in Michigan and the Fives Alive! Project in Ghana, which have been proclaimed as successful projects(19, 67).

These formal theories need to be complemented with informal theories. Stakeholder engagement to elicit underlying theories and identify and develop iterative context specific interventions has been acknowledged as an important step in theory based evaluation(35). Although these informal theories and local field experience make important contributions to the logic model, they are prone to selection bias. Given paucity of resources in health systems settings such as Malawi, stakeholder usually tend to focus their theories around structural changes such as new infrastructure, replenishing resources (such as drugs and supplies) and sustaining human resource, which might not be within the remit of an intervention project. In Malawi there is also a culture of *perdiemitis* that affects stakeholder engagement(234). While projects are geared towards achieving a higher social goal (safe motherhood, healthy newborns), individual incentives (per-diems, allowances) can take precedence over the larger social gain.

Moreover, most interventions are prescriptive, decided largely by donor and facilitating agencies before the beginning of the study. For instance MaiKhanda prescribed to IHI's Model for Improvement when other approaches to improvement were already in existence within country at the time of implementation. Since most projects are time bound and closely linked to the evaluation design, the degree of flexibility to incorporate informal theories to influence interventions is usually very limited.

Convincing stakeholders and getting their buy-in is crucial for intervention success. Evaluations that do not take informal theories into consideration have limited value in a theory-based evaluation. But on the other hand, if informal theories are to be part of the evolving intervention, then the evaluation design can be challenged by the intervention infidelity where mid-course correction of the intervention based on stakeholder feedback will lead to deviation from the original prescribed intervention. For instance, in the late intervention period it was felt that the intervention was not working as anticipated and a proposal to re-strategize the program and focus only on Salima district was suggested. This was based on the informal theories around the performance of health facilities in Salima. It was felt that there was a strong district leadership and the organizational 'readiness' was better in comparison with other districts (§6.7.2). This organizational readiness was characterized by the relatively better availability of human and material resources, better accessibility of the terrain and relatively better standing as women friendly facilities (§6.4.2.6). But a fixed evaluation design meant that the intervention could not be modified on the basis of the informal theories. On the other hand. even if evaluation design had been flexible and modified its design based on the emerging informal theories, the scope of current evaluation methods and tools to measure intervention fidelity and predict impact, is limited.

8.1.3 MaiKhanda's programme theory for QI

The program activities associated with MaiKhanda's improvement model were based on the theories of successful QI interventions elsewhere. Program

activities such as collaborative learning sessions in addition to knowledge sharing could in theory help build pressure within participating organizations to mimic other QI facilities. But the collaborative learning sessions were organized and lead mainly by MaiKhanda rather than senior programme leaders from within the healthcare system. This raised the issue of ownership and the intervention was perceived as something that was externally driven. While the mechanism of spread in the Michigan study was achieved through a ripple effect where the innovations in participating health facilities built pressure on other facilities to emulate them thus reaching a critical mass of QI facilities, at MaiKhanda participating facilities did not exhibit mimetic pressure on the other facilities. This is related to the issue of ownership mentioned above but could also be linked to the intervention's implementation strength.

There was no binding commitment for the health facilities to be involved in the QI activities since the decision to join the intervention was made at the MoH level. It was also difficult to get commitment from staff because of the frequent movement of staff within and across facilities. Although regular teleconferences and workshops provided a feedback mechanism, which lay the foundation for a networked community, their implementation was frequently hampered by logistics and financial challenges. MaiKhanda placed a lot of emphasis on data improvement exercises and feedback but the choice of key indicators such as NCFR was not a true reflection of the performance of the health facilities and anyways it was a complex measure to have warranted behaviour change at the individual or facility level. Checklists and elements of the change package provide a neutral platform for changing culture and practice around newborn care. But while checklists were introduced as part of the change packages, there was no organizational incentive for the Malawian health system to perform or change their culture of practice.

Program theories usually contain assumptions (either implicit or explicit) of how and why the program activities will bring a change, which are specific to each context and are subject to scrutiny within a TBE approach. Mechanisms,

which are an affirmation of the assumptions, cannot be directly observed-only inferred(220) and program theories help in drawing up this inference. Given the challenges on how to choose appropriate theories, it is likely that some of the mechanisms might have been overlooked.

8.1.4 MaiKhanda's implementation theory

Measuring implementation strength as an important precursor to intervention mechanism is not adequately acknowledged in literature. One of the reasons for using a modified theory based evaluation method for the MaiKhanda programme, combining Theory of Change and Realist Evaluation approaches, was to understand the link between intervention mechanism and implementation strength.

There are 2 major components to measuring implementation strength- it can be quantified in terms of the dose, duration, intensity and specificity of the intervention while implementation quality refers to the implementation fidelity.

Drawing an analogy from clinical studies, in terms of dosage of the intervention, MaiKhanda received a 'bolus' dose (through the QI Collaboratives) followed by 'sustained release' (equivalent to QI team meetings in the action period). This 'dose' was to be maintained at regular intervals both at the CEmOC level and health centre level, but this was not consistent and there were collaborative sessions conducted too close to each other followed by very short action periods affecting the process of continuous QI team engagement. The interventions got into full implementation mode as intended, only in the last 12-15 months of the intervention. Ovretveit(83), Alemi (225) and Pronovost(173) have identified the length of time required for successful QI Collaboratives to show outcomes is between 12 to 24 months. However this was much longer in Malawi. The intensity of the intervention was also very limited because of the limited capacity within MaiKhanda until the late intervention period. In terms of implementation specificity, the intervention did sway away from the agreed implementation plan, the focus being more on the CEmOC facilities and attention being paid to the neonatal change package only in the late intervention period.

Implementation theory can be plotted as a spider-web graph with the following elements, theoretical assumptions, intervention design, intervention implementation context and outcome, each scored on a likert scale (Figure 24). However, such a measurement is very subjective and not comparable with other interventions.

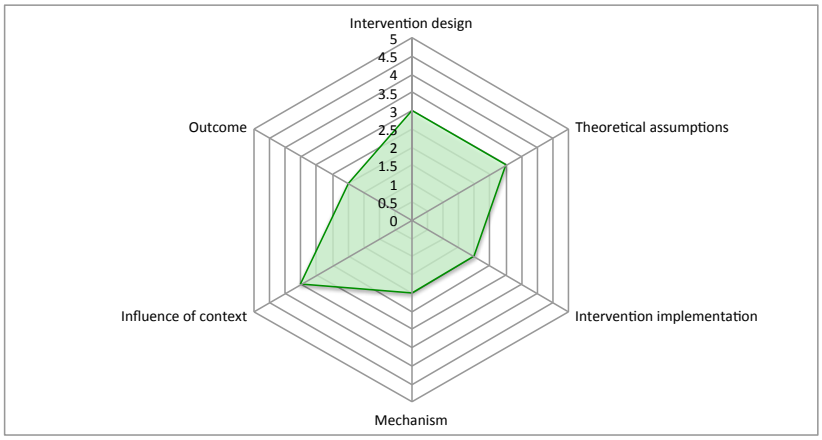

Figure 24: Measuring implementation (Illustrative only)

Plotting Implementation strength would include dose, duration, intensity, specificity and fidelity on a likert scale

One of the challenges for evaluation, in terms of implementation theory, is that there is no general consensus of a method for standardizing or objectively measuring individual elements of implementation and it is even more difficult to measure a pooled effect of the various elements.

But more importantly, it is difficult to determine how much implementation efforts are needed to achieve change in outcomes. Wilson et al have argued that while implementation effort is significantly associated with changes in activities related to the intervention, the implementation strength need not necessarily influence the outcome measures(101). This is because implementation strength cannot be measured in isolation as it is greatly influenced by the complexity of the intervention and of the context in which it is operating.

8.1.5 Complexity

I explore this assumption about the complexity of the intervention a bit further here. The basic premise of the model for improvement is for people to test their ideas within their local setting and where successful to share it with a larger peer forum through platforms such as the Collaborative learning sessions. Simple as the theory guiding QI intervention is, it is the actual design of the interventions that make it more complex.

As is true for most complex interventions, it is certain aspects of the intervention that make it complex rather than the intervention as a whole(235). For MaiKhanda, some elements of its intervention characteristics (such as intervention fidelity)(§6.3), its own internal context (§6.5), individual and organizational characteristics such as the subjective interpretation of the intervention by the human agency (§6.6) are factors that make it a complex intervention.

One of the major strategic changes in the early intervention period of MaiKhanda programme was to work across different levels of the health system so as to secure sustained and significant change. This change was influenced by The Health Foundation's (donor agency) learning and experience, which was shaping their own Theory of Change regarding quality improvement work. In essence, there was a recognition to work across different levels of the health system-the patient/community level, the clinical microsystem, the macro system (organization) and the environment, in order to make sustainable gains. In terms of intervention fidelity, this was different from the intervention design as envisaged in the original proposal. There the focus was on an integrated model for improvement combining criterion-based audits with rapid cycle improvements and use of MDAs as root cause analysis. Nevertheless, even in the revised Model for Improvement promoted by IHI there was a greater focus on improvement methods such as PDSAs and Collaborative cycles. However, a whole systems approach considering organizational maturity, staff motivation and MaiKhanda's internal stability and

capacity to embed interventions into existing health care systems perhaps would have been a better model for improvement, on hindsight.

Furthermore, MaiKhanda was 'facilitating' the intervention from outside the health system, with the actual implementation on the ground being conducted by the QI teams in the health facilities. This meant a long implementation chain with varying roles of the human agencies involved. The agencies involved donor perception, contextual understanding of technical advisors, the conceptual clarity of MaiKhanda facilitation team, the motivation of QI implementation team in the facilities, perception of staff in the facilities, the support of DHMT members in the district. In the process of introducing MaiKhanda interventions into the health system, the intervention had to manoeuvre itself through all these different agencies so as to effect implementation. MaiKhanda's degree of influence on the local teams would be determined by their level of engagement and degree of acceptability (through peer influence). While peer influence was a strong determinant of success in the Michigan study(67), MaiKhanda despite having QI officer with a clinical background in the late intervention period was considered external to the health system. The relationship between MaiKhanda staff and the QI teams on the ground could have been explored in greater detail as it bridges the important vital link between the NGO and the health system. Using methods such as social network analysis could help understand these vital links but also the power relation that influence this relationship (236).

Complexity also arises from a (health) system's inter-dependent parts. Because of this inter-dependence, improvement interventions need to focus across the different levels of the health system. A 'one-size fits all' methodology (such as the PDSA) might not be appropriate for improvement across other levels of the health system. By focussing on a single improvement method within just the microsystem level can make the gains from the improvement interventions unsustainable in the long run. For example: use of partograph in labour was affected by the non-availability of partograph forms; improvements in neonatal resuscitation was unsustainable

as a result of loss of institutional memory arising form the frequent transfer of nurses.

8.1.6 Context

Another important element influencing implementation strength is context and this includes both internal and external context. When MaiKhanda initiated the quality improvement initiative in Malawi, it was not being introduced in vacuum (or where nothing previously existed before). Even before the introduction of MaiKhanda interventions CBA (Criterion Based Audits) and IP (Infection Prevention) protocol and Performance assessment were being used in the Malawian health system. The Ministry of Health already had a technical working group on quality of care. This historical background and context is essential to understand when introducing QI concepts into the Malawian health care system. An early situational analysis done at the beginning of the intervention did describe this context. However, this was not taken into consideration in developing the intervention design and implementation. For example, it would have been easier for the QI intervention to gain acceptability into health facilities using CBA as the starting point, since this was an already established and accepted approach to improvement within the Malawian health care setting. Introducing a fairly new concept such as QI within existing models of quality of care, created conflict and confusion within the implementing agencies and with the health facility teams. This, perhaps, also contributed to the long lead-time in the programme interventions being adopted and implemented by the health facility teams.

MaiKhanda itself was an organization in transition. From starting off as a project, it later went on to register itself as a local entity in Malawi, with its own governance structure. This transition did consume a lot of the time of the senior management team and diverted their attention from the core components towards governance and institution building.

In MaiKhanda, there were isolated pockets of excellence within the QI intervention. These are the implementation theories contributing towards the common causal theory. But the impact of this (defined by facility NCFR) on

the overall programme was not seen because the critical mass to impact on the outcomes was perhaps not achieved. The programme would have benefitted if evaluation could have a greater focus into the implementation theories along side impact measurement of neonatal case fatality rate. I try to dwell deeper into the implementation theory here in my PhD, but this is done retrospectively and better evaluation designs could have been used, if the implementation theory was articulated at the beginning of the intervention. For instance, a case study approach of the 'outliers' identified in the control charts (for eg: best or worst performing facilities) could provide more insight into how interventions get accepted within the health system.

Other contextual factors that can influence implementation strength and it relation to outcomes are the extent of organizational readiness and organizational maturity within the health system. The Michigan study was able to reduce central venous line infection rates by applying QI to a complex clinical procedure among 108 hospitals within a span of 18 months(173). The degree of professional cohesion and organizational maturity that existed prior to start of the QI intervention could possibly have had a part to play in its success. But these elements were not described in the Keystone project nor were they adequately explored in the MaiKhanda project. Contextual 'software' factors such as organizational readiness and organizational culture influence implementation strength. There are iterative versions of measurement and classification of these 'softwares', such as organizational maturity index(107, 237) and COACH tool (238), but a consensus on standardization of these tools is required so as to help comparisons across interventions. Secondly, a close inter-relationship within different levels of the health system is key to intervention success at the microsystems level. In the Michigan study for example, there was a tight network between the hospital participating in the Collaboratives and the leadership within the hospital management was engaged in the improvement process(67). In contrast, in Malawi there is strong hierarchy between the higher levels of management (Ministry of Health officials) and the district and health facility staff (§6.6.5)

and this was generally seen as a hindrance for the cultivation of innovative ideas.

8.1.7 Intervention Characteristics

Closely related to the issue of context was that of the intervention characteristics. The MaiKhanda intervention is adopted largely from the Model for Improvement based on IHI's experiences in Northern America and Europe where QI interventions focussing specifically on the process of care have been successfully implemented. However, in comparison with the Malawian context these are relatively well functioning (resourceful) health systems. Introduction of QI concepts and techniques into less (or non) functional health systems needs to have a broader remit and needs to consider not just the care-giving processes but also the interdependent layers of service delivery mechanism and re-organization of care, in order to show impact. This concept of whole systems improvement approach is not new and has been suggested by others as well (23, 95), but the key challenge has been the implementation and evaluation of such an approach.

Another important intervention characteristic was that, in MaiKhanda, the interventions were mainly focussed on improvements in the process of care giving. For instance, the change package was mainly focussed on clinically excellent care at the microsystems (inpatient ward) level, conducting PDSA cycles on newborn resuscitation, management of sepsis and kangaroo mother care. Interventions focussing on the processes of care at the clinical microsystems level are able to generate localized improvement models, but without interacting with other layers across the health system, they are unable to develop a critical mass that will lead to measurable, sustainable and scalable improvements in outcomes.

8.1.8 Measuring implementation strength

Although context and complexity are important mediators of implementation strength, existing measurement methods for implementation strength do not take them into consideration(98). This is similar to a Theory of Change approach (in the evaluation of program theories) where context is considered

independent of the other components and evaluated separately (unlike in realist evaluation where context is an integral part of the intervention mechanism). The relationship between implementation strength and intervention outcomes is not linear and the 'tipping point' can only be achieved after a certain threshold of implementation strength is reached after which the change in outcome might be exponential. However, it is not possible to determine this threshold *a priori* and factor it in the evaluation design mainly because of its non-linear nature. It is also not clear if higher implementation strength will produce better outcomes. This will, to some extent, depend on how closely the various implementation elements reflect the theory of change.

In summary, the impact evaluation of QI intervention for reducing newborn mortality in 3 central districts in Malawi did not show an effect. This was probably related to its inadequate implementation strength, which did not carry enough momentum to trigger a change mechanism within the health facilities that MaiKhanda was working in.

8.2 Methodological issues

8.2.1 Theory Based Evaluation

There has been a plea and a thrust in recent years to look at wider research methodologies to evaluate complex interventions such as QI(169). For probability designs, in addition to RCTs, other approaches such as Bayesian methods (149) and Step Wedge Trials (SWTs) (172) have been suggested. In case of plausibility designs, theory based evaluation approaches have gained greater acceptability in recent years.

Theory based approach developed by Chen & Rossi in the 1980's is a paradigm shift from the methods based approach to evaluation (197). Data collection in theory based evaluation uses mixed methods and the mode of synthesis is qualitative and the approach follows a logic of configuration i.e. to assemble or arrange unlike conventional (or experimental) evaluation which follows the logic of aggregation. In conventional methods, strength lies in the numbers and the power to determine causation lies in effect size while in theory based evaluation, the focus is on context and mechanism to unearth the root cause of outcomes and causation is determined through the process of arranging critical pieces of evidence together. In conventional evaluation, the prevalence of cause-effect is more relevant than the mechanism of the intervention while in TBE, the mechanism behind the effect is more useful than the prevalence .

QI evaluation is interested in studying outliers and why they perform differently and TBE is well suited for evaluation of such outliers. Theory-based approaches can shed some light on how and why the intervention worked in certain setting and not in others.

TBE is an approach to evaluation that requires surfacing the assumptions that the program is based on, in considerable detail and following each step in the sequence to understand how the expected mini steps actually materialize. Theory based approach thus, by its very nature is multi-disciplinary and involves mixed methods. The main advantage is that TBE provides knowledge

of how or why a program is improving (or failing) thus providing leads on how to improve the existing program. There are various approaches to theory based evaluation (221) but Realist Evaluation and Theory of Change are the most commonly used applications of a theory-based approach(239).

The two commonly used approaches in TBE namely Theory of Change and Realist Evaluation have their strengths when it comes to implementation theory and program theory respectively and this is my rationale for using a hybrid TBE approach. While ToC hypothesizes links between programme activities and its anticipated outcome (and evaluates context independently), realist evaluation hypothesizes causal links between mechanisms released by an intervention and its anticipated outcome. Therefore theory testing in the former, focuses more on the implementation theory while in the latter theory testing is based on the program theory(198).

The hybrid theory based evaluation (TBE) approach used here in my PhD begins by mapping out the program theory, a rigorous evaluation of the context, mechanism and outcomes, followed by an analysis of the implementation theory.

8.2.2 Research Strategy

The use of research strategy, instead of a particular research method, consisting of a mix of methods to answer specific questions related to the program theory is the hallmark of theory-based evaluation of the MaiKhanda program (§4.1). A similar approach using comprehensive research strategy has been suggested by Greenhalgh and colleagues in their study on the modernization of health services in London, where they use an interpretive cases study design(217). While the MI (Modernization Initiative) study organized and collated primary data and produced preliminary thematic summaries, in MaiKhanda this was done using the CFIR framework. Both studies undertook evaluation at regular intervals but in the MI study, there was presenting, defending and interpreting particular action and events both within the research team and also stakeholders. These interpretations were further tested in subsequent rounds of evaluation and fragments of the case study

rewritten. In MaiKhanda similar meetings with key stakeholders existed through the data deep dive meetings and annual program management board meetings but this was mainly for information sharing and did not influence or modify the research strategy. The MI study used cross case comparisons to determine how the same mechanism played out in different context to produce different outcomes. There was a discussion at MaiKhanda about cross-comparisons, for instance focusing on a case study of Salima district. However it was not within the scope of research then to accommodate revisions to the evaluation design.

8.2.3 Comparing theories

A key question considered in my choice of research strategy is the methodological appropriateness to compare the MaiKhanda interventions with the successful QI interventions from the Keystone Project in Michigan, US. It can be argued that the 'context' is different in both the settings and that implementation was influenced by the particular context in which it was set, making the studies not suitable for comparison. While it is true that context does influence how the intervention eventually gets implemented, in terms of evaluation, the comparison was for the purposes of evaluating program theories (and the mechanisms associated with QI) and constantly building on them, not the individual activities, making such comparisons acceptable(164). An explanation offered by Pawson(202) stemming from Popper's philosophy, is to look at evaluation as a theory building exercise rather than a complete answer to the research question that can be generalized across different settings. Here program theories are considered repetitive and transferable and a source of cumulative evidence. By evaluating program theories of various similar (but unique) interventions, evaluation will eventually end up building up a picture of what works for whom and under what circumstance (and over what time duration). This more conventional approach to evaluation of program theories using realist evaluation approaches is through the application of a middle range theory (MRT). But this also has its limitations. It is prone to 'tunnel vision' of the intervention as the theories are defined by the

researcher and there are no set guidelines on how narrow or broad the MRT needs to be.

8.2.4 Data synthesis

I use the CFIR (Consolidated Framework for Implementation Research) (151) to synthesize the available data and organize the critical pieces of evidence together. The CFIR is very comprehensive and provides a wide range of domains and concepts to consider within a health systems intervention(151). One of the highlights of such a theory based programme evaluation is the use of all available forms of evidence, including the various evaluation studies, the annual and other reports developed by the implementing team. One limitation of the CFIR framework, though, is that it is not able to explain the interactions between the various elements. There are about 12 models specifically for implementation research, but there are no criteria or ranking of these models that would help researchers choose the most relevant framework for their evaluation(231). I chose CFIR as it was comprehensive but at the same time offered flexibility over the choice of constructs relevant to my research.

8.2.5 Program Theories

Program theories are at the crux of theory based evaluation. They ascertain the assumptions and help to draw inferences about the often unobservable or hidden intervention mechanisms. In realist terms, *"Interventions are theories incarnate"*(240)and realist evaluation considers a C-M-O configuration as a direct explanation of program theory. But in fact, interventions are theories in action and it is perhaps better to distinguish and explicitly outline the implementation of these interventions in the causal pathway. Thus a CIMO pattern configuration instead of CMO pattern configuration has been suggested (where 'I' stands for intervention implementation)(241). In terms of the C-M-O configuration itself, there is debate if it should be C→M=O or M→C=O i.e if it is the context that influence the mechanism or vice versa, to produce the outcome. Other views of the configuration define outcome as a function of the interaction between context and mechanism [O=*f* (M,C)] (242) while still others have considered outcome as an iterative interaction between

context and mechanism [O ←→ *f*(M,C)](241). My view is that context influences intervention implementation and intervention mechanism in different ways and that there is a constant interplay between implementation, context and mechanism to deliver the outcomes of an intervention. For an intervention to produce the desired outcomes certain mechanisms have to be 'fired' and this trigger is linked to the intervention's implementation strength, although not necessarily in a linear manner. Implementation plays a substantial role in shaping the context and mechanism of the intervention and vice versa and should be considered alongside programme theory. Consider the mechanism of 'harnessing data as a disciplinary force' (§7.1.5). MaiKhanda QI district teams, improved in the way of reporting which showed a maturity and understanding in what they were doing. The format of presentations, the language and the graphical presentation displayed an advanced level of understanding of the quality improvement in all the health centres [XXVIII]. This was approximately 36 months since data improvement workshops were being conducted as part of the intervention. Implementation factors such as data collection and learning sessions for data were consistently hampered by contextual factors such as shortage of human resources as well as the organizational context such as team maturity to view data as a learning point.

Evaluation of mechanisms in isolation using realist evaluation approach is perhaps best suited for evaluation of social programs but for interventions that are innovative pilot projects, yet to be embedded within a system, analysis of the implementation theory is perhaps a necessary precursor, before drawing inferences on the intervention mechanisms.

If there is good implementation data then mechanisms can be extracted directly from implementation theory. and then realist evaluation can be applied without necessarily having to develop a Middle Range Theory (MRT) (§ 4.3.4)

8.2.6 Implementation Theories

Implementation strength has been defined as *"The pooled effect of dose, duration, specificity and intensity of an intervention-in order to determine how much implementation efforts are needed to achieve a meaningful level of change in coverage and health outcomes."*(98) In addition to this quantitative estimation of a program strategy, it also incorporates some qualitative elements commonly referred to as fidelity of the intervention. A key drawback of current evaluation methods is that it does not differentiate between intervention fidelity and implementation fidelity.

The current practice in measuring implementation strength includes identifying essential components to be measured, grouping components into domains (thematic areas), building a measurement instrument and then piloting and measuring. Various methods have been suggested for measuring implementation strength such as use of different scales with detailed descriptions, percentage scales or binary scales showing if an intervention component has been implemented or not(98). One of the major drawbacks with the existing methods of measurement is that it treats each of the components as independent events and is not able to explain the inter-relationship between them. Nor is it clear if all elements have a common weightage. It also does not take the interaction effect with the context into account.

Another key limitation of current measurements in implementation strength is that it is viewed as a linear process. A host of factors come in and exit during various time periods of an intervention cycle all of which affect implementation strength differently.

Measurement of implementation strength tends to focus on outcomes rather than processes and is measured at definite time points rather than continuously(98). Alternative measurements such as the use of well annotated Statistical Process Control (SPC) charts as in the MaiKhanda intervention, is in my opinion a better tool for measuring implementation strength. Unfortunately, the SPCs in MaiKhanda was mostly focused on

measuring outcomes rather than processes and was not considered in the original evaluation plan to be used as a tool for measuring implementation strength.

The current methods are all static in nature i.e. they are measured at certain points in the intervention timeline whereas quality improvement is a continuous, dynamic, iterative process and needs to be measured as such. This is the reason that the implementation theory needs to be taken into consideration and not just implementation strength of the intervention. This is the approach that I have adopted in my study and I use the VICTORE framework to describe implementation strength as well as the associated factors influencing implementation strength such as context, emergence, length of implementation chain.

8.2.7 Criticism of Theory Based Evaluation approach

A recent systematic review of TBE approaches concluded that there is *"no evidence to support contention that theory based evaluations are more or less proper, useful, feasible or accurate than other forms of evaluation"*(243). According to some critiques, it is difficult to apply a theory-based approach to evaluation of complex adaptive systems that are dynamic in nature and unpredictable (244). Health systems are recognised as complex adaptive systems comprising a network of adaptive agents that exhibit system properties such as feedback loops, self–organization and emergence(245). While there is agreement that evaluation needs to take account of these properties as well as all the complex pathways(246, 247)and come up with a coherent explanation for the intervention, the use of TBE approaches to evaluate these complexity dimensions has been contested. Critiques argue that such complexity stretches the application of TBE to such a point where it becomes both methodologically and theoretically fragile(248). While others have argued that the basic tenet of evaluation designs using theory-based approaches is essentially to take account of this composite nature of theories in complex health systems and provide a meaningful understanding of the intervention(249).

One of the reasons for the methodological and theoretical weakness is interaction between different interdependent levels of health system and therefore mmultiple theories are needed to understand multiple levels of the system(194). Westhorp has suggested explaining these complex processes of change through layering of theories, where theories are organized within a hierarchy to reflect multiple levels of the system within the intervention. This has been described as “layering theories” (104). In the case of MaiKhanda, there could be the theory of profound knowledge as the overarching theory with diffusion of innovation theory explaining intervention scale up and spread, network and organizational theories to understand collective behaviour, normalization process theory to understand intervention embeddedness (and therefore sustainability). The list of theories to choose from is overwhelming(231) and logic models have been known to help refine the choice of theories.

An alternate approach would be to develop a middle range theory (MRT) as is usually done within a realist evaluation framework. A MRT is an intermediate between general theories of social systems and the specific intervention strategies(250). However, in using this approach there is a tendency to develop a ‘tunnel vision’ of the intervention mechanism, since the theories are developed *a priori* and specific to the intervention. However, the theories are formulated by the researchers (in consultation with the stakeholders rather than by the stakeholders) and are therefore prone to researcher bias.

One of the main criticisms of existing impact evaluation designs such as cRCT is that generalizability or external validity of an intervention is restricted. However, this issue is not resolved using a TBE approach either. Issue of generalizability still persist.

There has been criticism that TBE tends to prioritise program theories over the program itself(251). However, eventually, this prioritization and choice of evaluation design depends on the views of the funder and key stakeholders on what research questions they want answered.

Finally there is argument on whether the research should be prospective or retrospective. It is also not clear if having outcomes known (or not known) before drawing in on the theories will introduce some kind of a bias in theorizing what we know worked (or did not work)(252)

8.3 Key learnings

Evaluation of the MaiKhanda programme has been challenging from the inception for want of a proper definition of quality improvement. Quite often there is confusion regarding quality of care and quality improvement(142)

Quality of care can be thought of as an end product while quality improvement is a process that involves improving the quality of service delivery-which it is assumed will then improve quality of care. Quality improvement within a healthcare context can be defined in terms of improving quality of:

Processes of care: This is usually at the clinic or ward level and also sometimes referred to as the micro level. This involves the actual process of care giving such as 'dorikas' to escort patients to the labour room as in Kasungu district, use of partograph during labour. These processes of care are directly related to the knowledge and skills of the service provider, as well as their attitude, behaviour and belief systems. However, this process of care-giving is greatly influenced by the service delivery mechanism.

Service delivery mechanisms: This includes all the ancillary service linked to improving the process of clinical care at the ward level (micro level). This would include for instance resource availability, improving signal functions, structural relationship associated with human resources.

(Re)organization of healthcare systems: This includes the broader macro level such as the organization and the functional relationships associated with the different departments, human resources, understanding and balancing power relationships, engaging senior leadership.

In practice, focus is generally on the microsystem to improve processes of care. At MaiKhanda, focus was not even on the processes of care, rather it

was on the outcomes of care at the micro-systems level (i.e neonatal case fatality rate). Focus on outcomes can be misleading especially if one measures it at frequent time intervals.

The microsystem is where suffering actually gets relieved and everything else is ancillary to the microsystem to achieve its objective(95). But these different levels of the health system are interdependent on each other and greatly influence the microsystem.

8.3.1 Adopting a whole systems improvement approach

Literature on the success of QI is mainly from the developed countries. The context of the health system in these countries is very different to the set up that exists in developing countries. In Malawi, it is not just the human resources which is a challenge but most of the health system building blocks that are in various stages of incompleteness(59). QI operates well in building the inter-relationship between the various building blocks of the health system and therefore struggles in the absence of these building blocks. A certain minimum threshold of organizational readiness is necessary for QI to function well at the micro-system level. However, it is difficult to predict what minimum level of organizational readiness is needed before QI can kick in.

Improvement models, especially in developing countries such as Malawi need to take a whole systems approach to improvement (23, 95). This includes thinking not only about methods and tools such as PDSA cycles, but also considering other approaches, which are fit for purpose for different levels of the health system.

8.3.2 Institutional v/s Organizational Challenges

More often institutional challenges are mistaken as organizational challenges. Shortage of human resources is an institutional challenge, which cannot be resolved by individual QI teams developing PDSA cycles. This has to be approached as an industry problem. Shortage of health care workers globally is 7.2 million(253). In Sub-Saharan Africa, 11 of the 47 countries have no medical school for training and education in medical sciences. In Americas,

where 70% of the countries do have adequate health care workers, challenges even for basic universal health coverage exist in terms of distribution of the professionals, migration and appropriate training and skills mix. These are institutional challenges, which need to be differentiated from organizational challenges (179). The strategies to resolve them need a higher level of engagement across different levels of the health system.

Meanwhile fragmented structuring of the health care service delivery poses an organizational challenge. In Malawi, while newborn health is within the purview of the Reproductive Health Unit at the Ministry of Health, newborn illnesses fall within the remit of Integrated Management of Childhood Illnesses (IMCI) Unit. It is difficult to maintain continuity and quality of care in such a set up and is usually beyond the scope of most QI interventions to influence any change. Nevertheless, this needs to be taken into consideration during the design phase of the intervention.

The distinction between institutional and organizational challenge is important to keep in mind in the design of QI interventions as there will be certain factors beyond the scope of implementing QI teams and unrealistic expectations can hamper the morale of the team.

8.3.3 Intervention Design

The primary function of evaluation design is to be able to tell if and how the observed effects are indeed due to the intervention under consideration. Conventional impact evaluation methods have their limitations when evaluating complex interventions such as quality improvement. While it is necessary to understand if an intervention works or not, an equally important question to answer is 'what works for whom under what circumstances'. A plausibility design is perhaps best suited to answer this latter question and uses very different methods and tools from a probability design that is mainly used in impact evaluation to answer the former question(195). Evaluation needs to adopt a research strategy, rather than a specific research method, to be able to answer the evaluation questions.

Given the challenges of intervention and systems complexity and intervention fidelity, there is a general consensus in the research community to consider plausibility designs for evaluation alongside probability designs(195). They complement each other. Probability designs tell us 'if' an intervention works or not, plausibility attempts to answer the 'how' question. In other words, probability designs focus on the measurement of intervention outcomes while plausibility designs focus on the intervention mechanisms. Although there is consensus in principal, it is not clear how to operationalize such an evaluation.

One of the current debates in literature is whether RCT can be considered alongside realist evaluation to evaluate complex interventions. The debate captures the epistemological standpoints of RCTs and realist evaluations. It also highlights some of the operational challenges such as measuring context, capturing the non-linearity of complex interventions and inter-action of the human agency, and the choice of comparison group(28, 254, 255).

Comparison groups are usually a feature of probability designs and help to measure attribution of the QI intervention to the outcome. The choice of comparison group is therefore very important and is determined by the level at which the intervention takes place as well as the level at which outcomes are measured. QI interventions can focus at different levels such as improving the processes of care or service delivery mechanism or re-organization of the healthcare systems. In some cases it is a combination of these different levels. Furthermore, these different levels are inter-dependent thus making it difficult to isolate intervention effects even in the presence of a comparison group. Having a comparison group at a higher level such as district or organizational level, raises a challenge for adequate sample size for a quantitative analysis. While having a comparator cluster at a much lower level such as individual service provider level, raises challenges for appropriate comparative sample for outcome measurement.

There are challenges to selection of comparison groups at a conceptual level as well. In quality improvement interventions, the selection of groups or clusters (typically these are facilities within a health system) is influenced by

their level of commitment or “readiness” to be part of the intervention. Such facilities are bound to possess individual and organizational characteristics that are inherently different from their comparison groups. Moreover, evaluation designs using a comparison group require them to be selected at the beginning of the intervention. The diffusion of innovation approach might be a more organic and perhaps pragmatic approach from an implementation perspective, but this can be in conflict with most traditional evaluation designs where the selection of a comparison group happens *a priori*.

From an evaluation perspective, focusing on only the ‘low hanging fruits’ (innovators) does not tell us if the intervention has the same effect on the ‘not-so-enthusiastic’ (late adapters and laggards) facilities. A comparison group gives the opportunity to analyse and understand the mechanisms in play with ‘innovators and early adaptors’ as well as the ‘laggards’. However, from an improvement science perspective(46) psychology plays a role in the organic spread of QI interventions. Focusing on the low hanging fruit and managing to get a ‘critical mass’ of facilities that will adhere to QI principles, is a strategy that is more likely to influence the intervention acceptance by ‘later adapters & laggards’. This aspect of psychology of influencing a larger stakeholder group by building a ‘critical mass’ to eventually influence the outcome, is difficult to measure using a comparison group. It gets further complicated if the comparison groups have to be randomly allocated so as to control for measureable and unmeasurable confounding factors.

For impact evaluations, an important factor influencing evaluation design of a QI intervention is the indicator used to measure outcome and impact. Mortality measures alone are perhaps not the best indicator to measure intervention effect, given the relatively short period of time available for implementation (usually up to 5 years), the time lag between intervention implementation and effect and other contextual factors that influence the landscape for MNH. Evaluating performance measures such as the provision, utilization and coverage of interventions is equally important. Coverage especially is an important indicator of the interface between service delivery

(the managerial process) and the population (the epidemiological perspective). It can also be considered as a direct response to the implementation strength (dose)(256). Thus multiple outcomes need to be considered in the evaluation of whole system based improvement interventions.

A combination of probability and plausibility design is often misinterpreted as a mixed methods design containing qualitative and quantitative research methods. While probability designs are usually quantitative, plausibility designs could be qualitative or quantitative. For instance plausibility design can be a step-wedge trial with simulation model(257) which are purely quantitative evaluation methods.

Such a complex design requires a research strategy rather than a specific research method, which is another shift from a conventional evaluation design. The ultimate objective of such evaluation designs is about cumulative knowledge building by adding new knowledge to current understanding of programmes (202).

Improvement research is steadily gaining prominence within social and public health interventions. How interventions are going to be adapted, adopted and diffused within existing health systems should be an integral part of the evaluation rather than a hypothesis or a statement on sustainability of the intervention in the concluding remarks of a paper. This is because there are well-established evidence based interventions available, for the improvement of maternal and newborn health. As we move from Millennium Development Goals to Sustainable Development Goals, the challenge for practitioners and evaluators of health systems improvement is to see how interventions can be embedded within existing health systems and quality of care sustained(258). A wider health systems research approach is needed where implementation is considered as an organizational, social and political process involving stakeholder perceptions and their inter-relationships rather than being solely focused on service delivery(206). The challenge for systems improvement research is to capture this broad research agenda and structure it in a manner

that makes it pragmatic and operational and allows results to be used by implementers and policy makers alike.

8.4 Furthering the science

Current methods and practices in improving health systems research needs introspection. According to WHO, a main challenge to current health systems research is that health systems is still thinking in terms of its health service silos(259). Secondly, there is an assumption about the know-do gap being a linear process i.e when an evidence based clinical practice si presented there will e a rapid uptake of that practice.. But that is not so. In advancing the application of systems thinking in health, there is a need to move from vertical, compartmentalized thinking of health services models to systems thinking model that will help narrow the know-do gap.

Health systems thinking as suggested by WHO have 4 major steps for intervention design which includes convening stakeholders- collectively brainstorming- conceptualizing the effects and adapting and redesigning(260). However this is not a process that can happen at frequent intervals nor can the redesign be adapted very quickly at the grassroots level.

On the other hand, this process is very similar to the PDSA (Plan-Do-Study Act) cycle recommended in Quality Improvement, except that the suggested WHO model happens at a much higher level within the health system whereas PDSA cycles are conducted at the clinical microsystems level. Interventions at a higher level within the health system can address some of the health systems building block issues, but they can seldom effect the day to day routine work at the health facility level. For instance, data dashboards targeted at improving processes of care, are of little help if the intervention keeps focusing on case fatality rates in the health facilities. More proximal process indicators such as blood availability, use of partograph are more helpful from a grassroots systems improvement perspective. By focusing on small test of change at the ward level and then taking it to scale among peer groups, the intervention can build a critical mass to influence a whole systems

improvement. But this will still need the intervention design to consider the interdependent layers within the health system.

Current research in health systems is focused on investigating the individual components of the health systems building blocks, and there is very little research looking at system interdependence. When health systems component interact with each other, they exhibit systems properties such as reinforcing loops, non-linearity, disproportionate relationships (ie the critical mass required for change may vary in different settings) and emergence(248). These characteristics are difficult to measure and account for, using conventional research methods.

An alternate approach to research in health systems is needed which begins by identifying the limitations of current health systems research approaches and figures out new ways to evaluate the dynamic and evolving health systems. I suggest systems improvement research as an alternative. Some of the basic differences are outlined below (Table 13).

Table 13: Comparing Health Systems Research with Systems Improvement Research

	Health Systems Research	Systems Improvement Research
Target Audience	Policy Makers	People engaged in improving the system, of which policy makers are also a contributory force.
Centrality	Policy centered	Mechanism centered
Research Strategy	Independent/isolated mixed methods	Interlinked/inter-dependent mixed methods
	Focus on policy to improve how things are done at the grassroots level	Focus on change ideas at the grassroot level

Centrality matters. In health systems, research is geared towards the policy makers to engage in research and make change. This is a more top-down approach. The centrality in systems improvement research is towards improving service delivery mechanism and processes at the point of care. Governance and accountability mechanisms, across different levels of the health system, are geared towards supporting this process. Systems

improvement research thus focuses on understanding bottom up accountability mechanisms. The target audience in systems improvement research are mainly the people engaged directly in improving the system. This also includes, but not only, the policy makers among others.

The scope for systems improvement research is to evaluate the individual system components using conventional research methods but also use new and innovative research strategies to evaluate interdependence of the health systems components and its influence on interventions. This could use a research strategy similar to the one I have used in this study.

For other component of the intervention implementation framework, further research is needed to consider ways to measure organizational culture and context, individual beliefs, trust and accountability mechanisms. This can be achieved by adopting plausibility design such as TBE. Such approaches provide an understanding of the role of context and complexity in evaluating QI interventions. They are extremely useful for people seeking to implement similar programs but they are more like a well-informed hypothesis that needs to be tested in future evaluation. Alternative evaluation methods which combines TBE approaches with other impact evaluation designs such as Bayesian methods, step wedge trial with modelling and time series regression analysis have been suggested(149, 172, 261). These methods have the potential to incorporate research methods used in complex systems that have developed significantly over the past 10 years.

There is need to develop dialogue between health systems researchers and complex system modellers. Health system can be considered as a network of networks characterized by WHO building blocks but also of network dynamics i.e. how the different components of the system interact with each other thereby influencing intervention implementation, sustainability and scale up.

Such complex evaluation designs are best developed using a research strategy than having a single dominant research method. An ideal research strategy would include the following:

- Developing a program theory with key stakeholders. This involves identifying both formal and informal theories.
- Mapping out the causal chain: This can be done using a causal loop diagram to understand the actual process pathway for proposed intervention. This will give an idea about intervention fidelity as well as help uncover some of the implicit assumptions regarding the intervention
- Developing a probability design for an impact evaluation measuring improvements in health outcomes as well as intervention coverage.
- Understanding the context in which the intervention will operate.
    - A formative qualitative research to understand stakeholder perceptions.
    - A network analysis to understand the relationship between agencies at different levels within the health system.
- The next step would be collating all the available evidence. This can be done using CFIR or a similar framework.
- TBE approaches such as Theory of Change or realist evaluation or a combination of approaches can then help analyse the program theory .
- Interventions should also measure their implementation strength as part of their evaluation of the implementation theory.
- Research strategy may include a component of economic analysis depending on funding availability and the research objectives.

8.5 Limitations of the study

As my PhD was based on the secondary analysis of the data from MaiKhanda's comprehensive evaluation, limitations within that evaluation had implications for the availability and use of evidence within the theory-based evaluation.

MaiKhanda evaluation was not based on the research strategy outlined above. Infact, the evaluation's large focus was on the impact evaluation and the RCT design. While the process evaluation analysed the key areas of quality of care, it did not measure the dynamic relation between the different levels of the health systems and how it impacted on the quality of care provided.

The evaluation design for MaiKhanda programme was comprehensive but it had certain limitations. Key elements of the process evaluation such as tools for the staff motivation survey, provider knowledge survey were developed based on observations from the ground rather than using validated measurement tools available in literature. This limited the comparability of the data when used as a secondary source for my PhD.

There were not appropriate weightage applied to different variables within the staff psychology survey and provider knowledge survey and this could have influenced the results of the study. An exploration of the literature to standardize the variables was not done. Some of the studies such as the provider knowledge survey were only conducted at a single time point (rather than a continuous measure or at different time points) and therefore could not be compared over a period of time. Although there are objective tools available in literature to measure context and organizational culture, this was not used in the study.

While I do consider some of elements of complexity such as non-linearity, emergence and tipping point in my analysis, this was not included in MaiKhanda's original evaluation design and therefore data available was piece-meal and had to be inferred from the available data. .

There was limited monitoring data available. There was an assumption that the monitoring data would be readily available from the implementation team but this was not the case, limiting the analysis for theory-based evaluation. But regular programme updates were available through the biannual progress reports and the reflective evaluation that was conducted between the

evaluators and implementers towards the end of the project in 2010, which supported the development of the theory-based evaluation.

Since I was evaluating retrospectively, the selection of the various constructs for data synthesis and configuration was guided largely by the availability of data at hand rather than the ones most valuable to the study. For instance, the role of leadership was not independently investigated throughout the project. While health systems leadership was generally weak throughout the length of the intervention, it affect on the intervention or its embeddedness within the health system remains largely unknown.

And finally, literature review was focused on traditional approaches to evaluating quality improvement. Challenges of a uniform definition for quality and outcomes limited the scope of the literature search.

Chapter 9 Conclusion

A recent WHO multi-country study on maternal and newborn health concluded that there was no relation between health outcomes and high coverage with essential interventions, to reduce mortality in health care facilities(262). According to Horton the missing ingredient in this relation is quality of care(6).

Quality of care has been an integral part of healthcare and clinical practice since the time Semmelweis in 1847 segregated obstetric care and introduced hand washing techniques thereby greatly reducing case fatalities in maternity ward(39). Since then quality has been a lot more in focus in healthcare especially after Donabedian shared his seminal work on quality of care in 1980s(263). Quality improvement in healthcare has adopted techniques mainly from the manufacturing industry and has been used widely in Europe and US. However, evidence of their success in healthcare is not very conclusive, especially from low and middle-income countries. There have been limited efforts to critically analyse the techniques used in quality improvement interventions. One of the main challenges in evaluating quality improvement is the complexity of the interventions themselves and the complex nature of the systems in which they are implemented. The evidence regarding quality improvement interventions for resource poor settings is generally lacking.

My PhD starts off with an examination of the completed MaiKhanda trial which looked at the effect of QI interventions and community women's groups on maternal and newborn mortality in 3 central districts in Malawi(24). The study detected a 22% reduction in newborn mortality in the combined facility and community intervention clusters and a 16% reduction in perinatal mortality in the community intervention clusters. The study had statistical power to detect changes in neonatal mortality in both interventions individually. However, it failed to show an effect on newborn mortality for the QI interventions. The aim of my PhD is to understand why improvement intervention undertaken by MaiKhanda for newborn care failed to show an effect. Absence of effect could be attributed to a failure of theory, a failure of implementation, an evaluation

failure or a combination of these. In MaiKhanda, the lack of effect can be attributed mainly to a failure to implement key intervention strategies to a critical mass that would initiate changes to the mechanisms for improving service delivery and quality of care in the health facilities.

My primary objective was to understand the mechanisms by which the QI interventions worked and explore the interaction between the various factors that mediate the effect on neonatal mortality as observed in the cluster randomized control trial.

My research strategy consisted of developing a Theory of change, consolidating and synthesizing all the available evidence using an appropriate framework, analysing the program and implementation theory using theory based approaches to evaluation.

The Model for Improvement begins by introducing small rapid tests of change at the local health facility level through QI teams, using package of interventions such as Kangaroo Mother Care, Neonatal resuscitation and sepsis management. While challenges in introducing the PDSA cycles within the Malawian health systems was articulated in the initial project proposal, the project did not address this in their intervention design.

The Model for Improvement used in MaiKhanda is built around Deming's improvement theory(46) and Roger's diffusion of innovation theory(181). This theory considers improvement as a product of subject matter knowledge and profound knowledge. Subject matter knowledge on essential and emergency newborn care was generally lacking among health care providers. Similarly, understanding variations within the health systems is an acquired skill. While IHI through MaiKhanda, provided ample opportunities for people to learn the QI model, in general, QI teams lacked capacity to collate data and analyse the variations between the health facilities. QI was a fairly new concept in Malawi and MaiKhanda's attempts to embed it within existing health system was limited by challenges of the health systems context, MaiKhanda's own

organizational transition and QI and clinical capacity of health care providers. These findings are discussed further below.

The study presents a unique approach to data synthesis for a comprehensive evaluation design using the Consolidate Framework for Implementation Research (CFIR)(151). The synthesis takes into consideration the various reports and documents cumulated through the life of the project and juxtaposes them with the process evaluation studies conducted during the same period. In doing so, it draws a picture of the intervention with a multi-dimensional perspective, which provides insights into the evolution of the project. The framework is very comprehensive covering 5 major domains and a range of constructs, all of which were not included in my study. The choice of constructs was based on the availability of data rather than prioritizing the key constructs to consider. This is one of the limitations to my study. A drawback of the framework is that it does not enumerate the interaction between the different identified constructs.

CFIR helps to produce structured and comprehensive data that is then used for analysing the program theory in relation to the intervention outcome. I do this by comparing MaiKhanda's intervention strategies claimed to release the mechanism of the intervention and compare it with the program theories of Michigan Keystone Project which used Collaborative methods to successfully reduce their central venous line blood stream infections in 106 participating ICUs. The argument for such a comparison is that while the interventions per se are very unique and specific to their context, the program theories underlying the use of Collaborative methods in both the interventions is the same and therefore comparable. Theories offer a higher level of abstraction that can be comparable across different settings(202). The key finding from analysis of the program theory is that similar intervention strategies that triggered the mechanisms in the Keystone Project failed to generate mechanisms in MaiKhanda project.

The main challenge for MaiKhanda was to simultaneously implement and sustain the various change packages it had introduced in the different

facilities. While there were isolated instances of successful intervention activities within MaiKhanda, it did not build enough momentum to generate mechanisms across a critical mass of the facilities that would eventually result in improved newborn outcomes. This can be attributed to the implementation strength, context and complexity of MaiKhanda's interventions. This is explored in detail in my PhD, using the implementation theory.

Implementation theory was based on diffusion theories where better performing facilities were to act as role models for other facilities to emulate. The evaluation design posed a hindrance to the theorized spread of the intervention was the. The cRCT design for measuring impact evaluation required a random allocation of the improvement facilities and this conflicted with innovation diffusion theories, which prescribed a gradual organic spread of the interventions by strategically engaging the innovators and early adapters.

Limitations of the evaluation design notwithstanding, the implementation strength characterized by the dose, duration, intensity and specificity of the intervention was sub-optimal. There was a belief that given the dynamic nature of the intervention, the spread would be very rapid and perhaps even 'contaminate' the control sites of the RCT design. Contrary to this, the intervention struggled to maintain a constant cohort of QI teams in the facilities and programme was intense only in the last 18 months of the project. In my dissertation, I provide a descriptive analysis of implementation strength and highlight some of the measurement challenges in quantifying implementation strength in the Discussion Chapter.

Implementation strength is not the only factor triggering an intervention mechanism. Implementation strength tells only part of the story and an understanding of the context in which the intervention is operating is necessary in order to attribute implementation strength to changes in outcome. Implementation strength also cannot be measured independent of the intervention complexity. For example, MaiKhanda struggled to show an effect of its interventions, despite having a long pre-intervention period to

refine its interventions, while the Michigan study produced results within 18 month period. This can be because of other factors related to intervention complexity such as the long implementation chain for intervention delivery, the subjective perception of the agency (QI team) regarding QI and emergent properties of the intervention as well as contextual factors such as organizational readiness, the health systems context, QI team capacity to deliver QI interventions and MaiKhanda's own internal capacity.

The human agency is at the heart of implementation and the intervention required a continuous and prolonged time and effort, than was anticipated, to engage and train the health facility QI teams on the improvement model.

One of the key factors affecting the uptake of strategies was MaiKhanda's positioning within the health system and the degree of influence it could exert. This has a significant role to play in country where projects are donor supported and perhaps also donor driven. MaiKhanda lacked an 'insider effect' to influence change-the QI leadership at MaiKhanda did not belong to the professional community of nurses and clinicians which formed the bulk of the QI teams. The change agents such as the MaiKhanda technical lead and District QI officers were external to the Malawian health systems and the champions and leaders of QI from within the health system had competing priorities to engage at the level of detail that the intervention required them to. MaiKhanda's own capacity to be a facilitator to the QI teams was limited. The period of the intervention also saw MaiKhanda going through a period of rapid organizational transition, which affected intervention implementation on the ground. Furthermore, MaiKhanda's own understanding of QI concepts was evolving gradually and this coupled with its long implementation chain, influenced the subjective understanding of the QI teams regarding QI concepts. Health facility staff also lacked the necessary skills and knowledge related to management of newborn health.

Despite these shortcomings, QI interventions were implemented in the selected health facilities. Limited resources within the health facilities meant that gains achieved in some aspects of the intervention could not be sustained

in the long run. External contextual factors such as fuel shortages contributed to poor implementation. Changes in policy such as government ban on TBAs, affected intervention uptake. It resulted in an increase in health facility deliveries, overwhelming the already under-resourced staff capacity in the health facilities. It is conceivable that quality improvement was not on top of their priority list.

But, motivation to be involved in QI Collaboratives, especially in resource constrained settings, can be influenced by the lure of personal incentives (such as per diems for attending workshops and meetings) as much as individual's commitment to broader social gains (ie reduction in newborn case fatality rates in their facility). Moreover, the improvement model was competing against other existing models and it was difficult to get enough stakeholder commitment to the prescribed model as there were huge expectations fuelled by the poverty and poor governance structures and a culture of "*perdiemitis*" was prevalent in Malawian health care system(234).

While the Model for Improvement tends to focus at the clinical microsystem level and have been successfully implemented in other settings, Malawi in comparison to these settings, had system level limitations of human and material resources as well as organizational capacity. Perhaps a whole systems improvement would have been more appropriate for MaiKhanda's interventions rather than a set of change packages focused at the clinical microsystems level only. A whole systems improvement framework considers health care as complex adaptive systems, which expands the scope of quality improvement definition thus considering the interaction between the different components of the intervention as well as its relation to the context(224).

9.1 TBE approaches to understand QI interventions

In evaluating QI interventions, the process of introducing the intervention into a health system needs further consideration mainly because of the human agency involved(183). Unlike clinical trials, the dosage in public health interventions is unpredictable-a small (or large) stimulus i.e implementation dose, can invoke a greater (or lesser) participation of the human agency, both

at an individual level and collectively, based on their belief in the intervention, self efficacy, their individual motivation and relationship with their organization and contextual factors such as organizational culture. As was observed in the case of MaiKhanda, individual self-efficacy and health system readiness were key factors affecting intervention uptake and embeddedness within the health system.

Impact evaluation methods have their limitations when evaluating whole systems improvement interventions and by themselves are not sufficient. It cannot not tell us how or why an intervention worked (or not). A paradigm shift from the methods based evaluation is theory based evaluation (TBE) approach suggested by Chen & Rossi in 1980's (197). For a TBE, interventions need to have established program theories (PT). The use of theories in improvement evaluation is under-recognized and under-utilized mainly because researchers are mystified with theories as being very abstract and irrelevant(264). Explicit or not, practitioners invariably do use theory in the development of their interventions. One of the first steps in TBE is to make these assumptions and theories more explicit.

Mechanisms are not a set of activities; rather they are underlying causal levers that bring about change in the reasoning and behaviour of individuals and organizations. We use theories to elicit these mechanisms, as mechanisms cannot be directly observed, but only inferred(220). It is likely that within the boundaries of established program theories for MaiKhanda project (i.e theory of profound knowledge and diffusion of innovation theory) I might have overlooked some of the mechanisms at play. For instance, consideration of social network theories and normalization process theory perhaps, could have helped to understand the mechanisms to navigate and embed interventions across different levels of the health system. The choice of theories plays an important part in evaluation using theory-based approaches.

Evaluating interventions as purely ‘theories incarnate’ (202) can be less meaningful if the implementation strength is not measured alongside it. This

is one of the limitations of a realist evaluation approach which looks at interventions purely in terms of its theories. However, measuring implementation strength itself is a challenge and it is further compounded by its interaction with context and influenced by intervention complexity. This is especially challenging for evaluation when interventions have defined and limited timelines.

9.2 Issues and challenges in evaluating quality improvement in resource poor health settings

Quality of care has been described as the third revolution in global health after the metrics revolution and the revolution in accountability(6). The first challenge with quality lies in defining it(17, 63, 142). Within the context of resource poor settings, quality improvement can be perceived as the optimal delivery of evidence based care by co-creating knowledge and re-organizing health systems to be resilient and responsive. This broadens the definition and scope of quality improvement intervention and challenges the status quo. The status quo is, quality improvement interventions are usually focused on a singular improvement technique or approach to resolve issues of quality across the various inter-dependent levels of the health system. However, there is a push to consider whole systems improvement approach, especially while applying QI in resource poor settings.

From an evaluation perspective, one of the main challenges is to encompass the whole systems improvement approach within a single evaluation framework and link it to health outcomes. Innovative impact evaluation designs such as step-wedge design(96) and simulation models(261) or Bayesian methods(149, 155) have been suggested.

However, given the challenges of intervention and systems complexity and intervention fidelity, a general agreement in evaluation is for plausibility designs to be considered alongside probability designs(27, 29, 171, 195).

A little explored area is the role of comparison groups in plausibility designs where there is a need to understand differences in context and mechanisms

between intervention and comparison sites. Quality improvement models are based on diffusion theories, which makes the spread of interventions organic and is dictated largely by interpersonal networks. Comparison groups can be considered as a hindrance to the organic spread of such QI interventions. The choice of comparison groups in such evaluation designs therefore, requires carful consideration.

Recently, realist RCTs have been proposed as an approach to evaluate complex interventions such as QI(28). In such trials, the choice of comparison groups is likely to be related to the key intervention processes and functions, rather than the precise activities itself(164). Such approaches to evaluation have to take into account, three C's: the context, the complexity and the choice of an appropriate comparison group.

As is evident from the study, a single research method will not be able to provide justice to evaluation of a complex set of factors that influence newborn outcomes. I propose a research strategy that includes developing a Theory of Change, followed by evaluation of the program theory, measuring implementation strength, analysing implementation theory and comparing this in relation to the outcomes of the intervention observed through the impact evaluation. The results arising from such a comprehensive evaluation will contribute to the growth of improvement science with the accumulation of knowledge and explanation rather than just a *bedrock of observational facts* (265).

9.3 Recommendations for policy and practice

Policy recommendations:

- A whole systems improvement approach is needed for addressing system level inter-dependencies in healthcare. Failure to address interdependencies is particularly damaging in low-income settings where, improvement interventions tend to fall into the trap of generating 'patchworks of excellence', with knock-on effect on other parts of the system. A systems improvement approach should place greater emphasis

on improving processes (along with outcomes) at different levels of the health system.

- Improvement science is a fairly recent concept in developing country health systems. Greater investments are required in implementation and improvement research in these settings.
- Capacity building is required in using the science of improvement as well as evaluating it. Building the capacity of health care providers and stakeholders has to be a key function of improvement interventions atleast in the early phase of the intervention.
- Skills building needs to be incorporated into the routine functioning of the health system, preferably through on-site training so as to avoid staff absenteeism from the facilities to attend training workshops.

Programme changes

- In implementing a QI intervention, simplicity of intervention is key. There needs to be a clear protocol to guide frontline workers. Projects need to consider a design that will be able to capture and document improvement without any extra burden on the health facility staff.
- Investment in additional support staff as well as technology is required especially in the initial period of the study, for data capture and for demonstrating the value in improving quality of care.
- Intervention design needs to factor in adequate time for improvement 'learning curve' when designing the intervention. Robust feedback loops, backed up by routine mentorship and tailored to practice for the individual or team is needed as a vehicle through which to implement and actualize QI.
- QI interventions should be closely aligned with the strategic goals of the organization as this increases their likelihood of success.
- Improvement ideas for testing need to be feasible and should ensure that the improvement methods such as PDSA cycles are completely implemented.

- Data collection is crucial for improvement and can be more efficient if linked with evaluation data collection.

Further Evidence

- The evidence base regarding the role of leadership and their attribution to the success of an intervention remains thin and needs further investigation.
- Improvement interventions can be human resource intense and some preliminary work and learning to understand organizational readiness is required before introducing a QI intervention in a healthcare setting.
- Innovative research designs are also required to understand the linkages between different levels of the health systems and how that improves quality at the point of care.
- QI emphasizes participation and engagement of local health facility staff from the design phase of the intervention. An understanding of organizational culture is needed to do this. Application of existing frameworks such as the competing values framework can provide a better understanding of the culture and practice existing in resource poor settings.
- Evaluation of complex intervention such as QI needs to look at multitudinous factors and the interaction between them. This can be achieved by developing a Theory of Change and building a research strategy around it.
- A research strategy should include evaluation of the context and measurement of implementation strength, using a range of data sources such as the district health information systems, project monitoring data, improvement data and evaluation data.
- Harmonizing this data and optimizing their use for evidence requires partner commitments as well as significant investments in improving data management.
- Analysing implementation strength is particularly important to explain the variances in quality of care across different settings. Measuring implementation strength needs to take the context and intervention complexity into consideration.

- A theory-based evaluation should include theorization of potential intervention harms when developing the ex-ante theory of change

Alternative approaches to innovative evaluation models for QI interventions need to be considered, such as adapted randomised-controlled trials and realist RCTs.

Bibliography

1. Kassebaum NJ, Bertozzi-Villa A, Coggeshall MS, Shackelford KA, Steiner C, Heuton KR, et al. Global, regional, and national levels and causes of maternal mortality during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2014;384(9947):980-1004.
2. World Health Organization. Trends in Materal Morality:1990 to 2013, Estimates by WHO, UNICEF, UNFPA, The World Bank and the United Nations Population division, Executive Summary. 2014;WHO/RHR/14.13.
3. Say L, Chou D, Gemmill A, Tuncalp O, Moller AB, Daniels J, et al. Global causes of maternal death: a WHO systematic analysis. The Lancet Global health. 2014;2(6):e323-33.
4. Liu L, Oza S, Hogan D, Perin J, Rudan I, Lawn JE, et al. Global, regional, and national causes of child mortality in 2000-13, with projections to inform post-2015 priorities: an updated systematic analysis. Lancet. 2015;385(9966):430-40.
5. Darmstadt GL, Bhutta ZA, Cousens S, Adam T, Walker N, de Bernis L, et al. Evidence-based, cost-effective interventions: how many newborn babies can we save? Lancet. 2005;365(9463):977-88.
6. Horton R. Offline: The third revolution in global health. The Lancet. 2014;383.
7. World Bank. Malawi GDP per capita, PPP (current international $) 2013. Available from: http://databank.worldbank.org/data/views/reports/chart.aspx.
8. National Statistical Office M. MDG Endline Survey-Key Findings 2014. Available from: http://www.unicef.org/malawi/MLW_resources_mes2014.pdf.
9. Ministry of Health M. Malawi National Health Accounts, with sub-accounts for HIV/AIDS, Malaria and Reproductive Health and Child Health for Financial Year 2009/10, 2010/11 and 2011/12. 2014.
10. Kanyuka M, Ndawala J, Mleme T, Chisesa L, Makwemba M, Amouzou A, et al. Malawi and Millenium Development Goal 4: Early Adopter, Early Achiever. Lancet Global Health (under peer review). 2015.
11. Kinney MV, Kerber KJ, Black RE, Cohen B, Nkrumah F, Coovadia H, et al. Sub-Saharan Africa's Mother, Newborns and Children: Where and Why do they die? PloS Med. 2010;7(6).
12. Zimba E, Kinney MV, Kachale F, Waltensperger KZ, Blencowe H, Colbourn T, et al. Newborn survival in Malawi: a decade of change and future implications. Health Policy Plan. 2012;27 Suppl 3:iii88-103.
13. Thaddeus S, Maine D. Too far to walk: maternal mortality in context. Soc Sci Med. 1994;38(8):1091-110.
14. Manandhar DS, Osrin D, Shrestha BP, Mesko N, Morrison J, Tumbahangphe KM, et al. Effect of a participatory intervention with women's groups on birth outcomes in Nepal: cluster-randomised controlled trial. Lancet. 2004;364(9438):970-9.
15. Lewycka S, Mwansambo C, Rosato M, Kazembe P, Phiri T, Mganga A, et al. Effect of women's groups and volunteer peer counselling on rates of mortality, morbidity, and health behaviours in mothers and children in rural Malawi (MaiMwana): a factorial, cluster-randomised controlled trial. Lancet. 2013;381(9879):1721-35.

16. Prost A, Colbourn T, Seward N, Azad K, Coomarasamy A, Copas A, et al. Women's groups practising participatory learning and action to improve maternal and newborn health in low-resource settings: a systematic review and meta-analysis. Lancet. 2013;381(9879):1736-46.
17. Pittrof R, Campbell O, Filippi V. What is quality in materntiy care? An international perspective. Acta Obstet Gynecol Scand. 2002;81:277-83.
18. Franco L, Marquez L. Effectiveness of collaborative improvement: evidence from 27 applications in 12 less-developed and middle income countries. British Medical Journal. 2011(20):658-65.
19. Nana A.Y Twum-Danso AG, Osafo E, et al. A nationwide quality improvement project to accelerate Ghana's progress toward Millenium Development Goal Four: design and implementation progress. International Journal for Quality in Health Care. 2012;24(6):1-11.
20. Schouten LM HMea. Evidence for the impact of quality improvement collaboratives: a systematic review. British Medical Journal. 2008;336(7569):1491-4.
21. . Available from: http://www.hciproject.org.
22. Boucar M FL, Sabou D, et al. Sustaining better maternal and newborn care and quality improvement in Niger: Challenges and successes. USAID Health Care Improvement Project-Research and Evaluation. 2011.
23. Donabedian A. The quality of care. How can it be assessed? JAMA. 1988;260(12):1743-8.
24. Colbourn T, Nambiar B, Bondo A, Makwenda C, Tsetekani E, Makonda-Ridley A, et al. Effects of quality improvement in health facilities and community mobilization through women's groups on maternal, neonatal and perinatal mortality in three districts of Malawi: MaiKhanda-a cluster randomized controlled effectiveness trial. International Health. 2013;5(3):180-95.
25. Wight. D, A O. Unpacking the 'black box': the importance of process data to explain outcomes. Stepehenson. J, Imrie.J, Bonnell.C, editors. Oxford: Oxford University Press; 2003.
26. Bonell CP, Bennett R, Oakley AO. Sexual health should be subject to experimental evaluation. In: Stephenson J, Imrie J, Bonell C, editors. Effective sexual health interventions: Issues in experimental evaluation. Oxford: Oxford University Press; 2003. p. 151-66.
27. Oakley A, Strange V, Bonell C, Allen E, Stephenson J, Team RS. Process evaluation in randomised controlled trials of complex interventions. Bmj. 2006;332(7538):413-6.
28. Bonell C, Fletcher A, Morton M, Lorenc T, Moore L. Realist randomised controlled trials: a new approach to evaluating complex public health interventions. Soc Sci Med. 2012;75(12):2299-306.
29. Strange V, Allen E, Oakley A, Bonell C, Johnson A, Stephenson J, et al. Integrating Process with Outcome Data in a Randomized Controlled Trial of Sex Education. Evaluation. 2006;12(3):330-52.
30. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. Bmj. 2000;321(7262):694-6.
31. Hulton A.L MZ, Stones R.W. A framework for the evalaution of quality of care in maternity services. University of Southampton. 2000.

32. Colbourn T, Pulkki-Brannstrom AM, Nambiar B, Kim S, Bondo A, Banda L, et al. Cost-effectiveness and affordability of community mobilisation through women's groups and quality improvement in health facilities (MaiKhanda trial) in Malawi. (1478-7547 (Electronic)).
33. Colbourn T, Nambiar B, Costello A. Evidence: MaiKhanda Final evaluation report. The Health Foundation. 2012:218.
34. Birckmayer JD, Weiss CH. Theory-based evaluation in practice. What do we learn? Evaluation review. 2000;24(4):407-31.
35. White H. Theory-Based Impact Evaluation. International Initiative for Impact Evaluation. 2009;Working Paper-3:1-20.
36. Slaney G. Foreword In Pollock A, Evans M. Surgical Audit. London, Buttersworth. 1989:v-vi.
37. Devlin H. Audit and the quality of clinical care. Ann R Coll Surg Eng. 1990;72:3-14.
38. Best M, Neuhauser D. Ignaz Semmelweis and the birth of infection control. Qual Saf Health Care. 2004;13:233-4.
39. Noakes TD, Borresen J, Hew-Butler T, Lambert MI, Jordaan E. Semmelweis and the aetiology of puerperal sepsis 160 years on: a historical review. Epidemiology and Infection. 2008;136(01):1-9.
40. Marie R Abraham SB, Diane M.Ashton, Maribeth Badura, Vani R.Bettegowda, Eric Beiber. Toward Improving Outcomes for PregnancyIII. March of Dimes.
41. Taylor F. Shop Management. New York: Harper and Brothers. 1919.
42. Shewhart WA. Economic control of quality of manufactured product. New York. 1931;501.
43. Shewhart WA. Quality Control Charts. Bell System Technical Journal. Unknown:593-603.
44. Best M, Neuhauser D. Walter A Shewhart, 1924, and the Hawthorne factory. Qual Saf Health Care. 2006;5:142-3.
45. Berwick DM. Controlling variation in healthcare: A consultation from Walter Shewhart. Medical Care. 1991;29(12):1212-25.
46. Langley M, Nolan, Nolan, Norman, Provost. The Improvement Guide, 2nd Edition. Josey-Bass. 2009.
47. Teich ST, Faddoul FF. Lean Management-The Journey from Toyota to Healthcare. Rambam Maimonides Med J. 2013;4(2).
48. The Health Foundation. Quality Improvement made simple. 2013.
49. Donabedian A. Evaluating the Quality of Medical Care. The Milbank Memorial Fund Quarterly. 1966;44(3):166-206.
50. Avedis D. Promoting Quality through evaluating the process of patient care. Medical Care. 1968;6(3):181-202.
51. John M.Luce ABB, Philip R.Lee. A brief history of health care quality assessment and improvement in the United States. West J Med. 1994;160:263-8.
52. Weiner BJ, Alexander JA, Shortell SM, Baker LC, Becker M, Geppert JJ. Quality Imporvement implementation and hospital performance on quality indicators. Health Serv Res. 2006;41(2):307-34.
53. Donabedian A. Promoting Quality through Evaluating the Process of Patient Care. Medical Care. 1968;6(3):181-202.
54. Berwick DM. Continuous improvement as an ideal in health care. N Engl J Med. 1989;320(1):53-6.

55. Sherman J, Malkmus M. Integrating Quality Assurance and Total Quality Management/Quality Improvement. JONA. 1994;24(3):37-41.
56. Banerjee A, Stanton E, Lemer C, Marshall M. What can quality improvement learn from evidence-based medicine? Journal of the Royal Society of Medicine. 2012;105(2):55-9.
57. Bridgewater B, Irvine D, Keogh B. NHS transparency. Bmj. 2013;347:f4402.
58. Liu C, Babigumira J, Chiunda A, Katamba A, Litvak I, Miller L, et al. Finding the best examples of healthcare quality improvement in Sub-Saharan Africa. Qual Saf Health Care. 2010;19(5):416-9.
59. Smits HL, Leatherman S, Berwick DM. Quality improvement in the developing world. International journal for quality in health care : journal of the International Society for Quality in Health Care / ISQua. 2002;14(6):439-40.
60. Lee RI, Jones LW. The Fundamentals of good medical care. University of Chicago Press, Chicago. 1933.
61. Donabedian A. Explorations in Quality Assessment and Monitoring, . Health Administration Press, Ann Arbor MI. 1982;2.
62. Raven JH, Tolhurst RJ, Tang S, van den Broek N. What is quality in maternal and neonatal health care? Midwifery. 2011.
63. Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century: The National Academies Press; 2001.
64. World Health Organization. Quality of care: a process for making strategic choices in health systems. 2006.
65. Marshall M. Applying quality improvement approaches to health care. BMJ Qual Saf. 2009;339(b3411):819-20.
66. Marshall M, Pronovost P, Dixon-Woods M. Promotion of improvement as a science. Lancet. 2013;381(9864):419-21.
67. Dixon-Woods M, Bosk CL, Aveling EL, Goeschel CA, Pronovost P. Explaining Michigan: Developing an Ex Post Theory of a Quality Improvement Program. The Milbank Quarterly. 2011;89(2):167-205.
68. Batalden PB, Davidoff F. What is "quality improvement" and how can it transform healthcare? Qual Saf Health Care. 2007;16(1):2-3.
69. Perla RJ, Parry GJ. The epistemology of quality improvement: it's all Greek. BMJ Qual Saf. 2011;20 Suppl 1:i24-7.
70. Thor J, Lundberg J, Ask J, Olsson J, Carli C, Harenstam KP, et al. Application of statistical process control in healthcare improvement: systematic review. Qual Saf Health Care. 2007;16(5):387-99.
71. Perla RJ, Provost LP, Parry GJ. Seven propositions of the science of improvement: exploring foundations. Quality management in health care. 2013;22(3):170-86.
72. Benneyan J. Statistical Quality Control methods in infection control and hospital epidemiology, Part II: Chart use, statistical properties and research issues. Infect Control Hosp Epidemiol. 1998;19(4):265-83.
73. Benneyan J, Lloyd R, Plsek P. Statistical process control as a tool for research and healthcare improvement. BMJ Qual Saf Health Care. 2003;12:458-64.
74. Marshall M, Mountford J. Developing a science of improvement. Journal of the Royal Society of Medicine. 2013;106(2):45-50.

75. Glasziou P, Ogrinc G, Goodman S. Can evidence-based medicine and clinical quality improvement learn from each other? BMJ Qual Saf. 2011;20 Suppl 1:i13-7.
76. Nolan TW. Execution of Strategic Improvement Initiatives to produce system level results. IHI Innovation Series White Paper. 2007.
77. AE Powell RR, HTO Davies. A systematic narrative review of quality improvement models in healthcare. 2009:167.
78. FIllingham D. Canlean save lives? Leadership in Health Services. 2007;20(4):231-41.
79. Kilo CM. A framework for collaborative improvement: lessons from the Institute for Healthcare Improvement's Breakthrough Series. Quality management in health care. 1998;6(4):1-13.
80. S.Mittman B. Creating the evidence base for Quality Improvement Collaboratives. Ann Intern Med. 2004;140(11):897-901.
81. Geboers H, van der Horst M, Mokkink H, van Montfort P, van den Bosch W, van den Hoogen H, et al. Setting up improvement projects in small scale primary care practices: feasibility of a model for continuous quality improvement. Quality in health care : QHC. 1999;8(1):36-42.
82. Berwick DM. Developing and testing changes in delivery of care. Ann Intern Med. 1998;128(8):651-6.
83. Ovretveit J, Bate P, Cleary P, Cretin S, Gustafson D, McInnes K, et al. Quality collaboratives: lessons from research. Qual Saf Health Care. 2002;11(4):345-51.
84. Greenhalgh T, Robert, G., Bate, P., Kyriakidou, O., Mcfarlane, F. and Peacock, R. How to spread good ideas: a systematic review of the literature on diffusion, dissemination and sustainability of innovations in health service delivery and organization. NCCSDO, London. 2004.
85. Kosseff AL, Niemeier S. SSM Health Care clinical collaboratives: improving the value of patient care in a health care system. The Joint Commission journal on quality improvement. 2001;27(1):5-19.
86. Bate P, Robert, G. and McLeod H. Report on the '*Breakthrough'* Collaborative approach to quality and service improvement within four regions of the NHS. University of Birmingham, Birmingham. 2002.
87. Savage J, Scott C. The modern matron: a hybrid management role with implications for continuous quality improvement. Journal of Nursing Management. 2004;12:419-26.
88. Walley P, Gowland B. Completing the circle: from PD to PDSA. International Journal of Health Care Quality Assurance. 2004;17(6):349-58.
89. Leatherman S. Optimizing quality collaboratives. Qual Saf Health Care. 2002;11(4):307.
90. Newton PJ, Davidson PM, Halcomb EJ, Denniss AR, Westgarth F. An introduction to the collaborative methodology and its potential use for the management of heart failure. The Journal of cardiovascular nursing. 2006;21(3):161-8.
91. Malenka D, O'Connor G. The Northern New England cardiovascula disease study group. The Joint Commission journal on quality improvement. 1998;24(10):594-600.
92. Horbar J, Rogowski J, Plesk P. Collaborative quality improvement for neonatal intensive care. Pediatrics. 2001;107:14-22.

93. Hulscher ME, Schouten LM, Grol R. Collaboratives. The Health Foundation. 2009.
94. Debra dS. Improvement Collaboratives in health care. The Health Foundation, Evidence Scan. 2014.
95. Berwick DM. The John Eisenberg Lecture: Health Services Research as a Citizen in Improvement. Health Serv Res. 2005;40(2):317-35.
96. Speroff T, Gerald TOC. Study Designs for PDSA Quality Improvement Research. Q Manage Health Care. 2004;13(1):17-32.
97. Marshall M. Seeing with new eyes: what can general practice learn from the science of improvement? The British journal of general practice : the journal of the Royal College of General Practitioners. 2011;61(586):352-5.
98. Schellenberg J, Bobrova N, Avan B. Measuring implementation strength: Literature review draft report 2012. 2012.
99. Provost LP. Analytical studies:a framework for quality improvement design and analysis. BMJ Qual Saf. 2011;20 (Suppl1):i92-i6.
100. Hulscher ME, Schouten LM, Grol RP, Buchan H. Determinants of success of quality improvement collaboratives: what does the literature show? BMJ Qual Saf. 2013;22(1):19-31.
101. Wilson MG, Basta TB, Bynum BH, DeJoy DM, Vandenberg RJ, Dishman RK. Do intervention fidelity and dose influence outcomes? Results from the move to improve worksite physical activity program. Health Educ Res. 2010;25(2):294-305.
102. Lilford RJ, Brown CA, Nicholl J. Use of process measures to monitor the quality of clinical practice. Bmj. 2007;335(7621):648-50.
103. Stetler C, Legro M, Wallace C, Bowman C, Guihan M, Hagedonr H, et al. The role of fomrative evaluation in implementation research and the QUERI experience. J Gen Intern Med. 2006;21(Suppl 2):S1-8.
104. Westhorp G. Developing complexity-consistent theory in a realist investigation. Evaluation. 2013;19(4):364-82.
105. Scott T, Mannion R, Davies H, Marshall M. The quantitative measurement of organizational culture in health care: a review of the available instruments. Health Serv Res. 2003;38(3):923-45.
106. Kaplan HC, Brady PW, Dritz MC, Hooper DK, Linam WM, Froehle CM, et al. The influence of context on quality improvement success in health care: a systematic review of the literature. Milbank Q. 2010;88(4):500-59.
107. Lombarts MJ, Rupp I, Vallejo P, Sunol R, Klazinga NS. Application of quality improvement strategies in 389 European hospitals: results of the MARQuIS project. Qual Saf Health Care. 2009;18 Suppl 1:i28-37.
108. Scott T, Mannion R, Marshall M, Davies H. Does organisational culture influence health care performance? A review of the evidence. J Health Serv Res Policy. 2003;8(2):105-17.
109. Kaplan HC, Provost LP, Froehle CM, Margolis PA. The Model for Understanding Success in Quality (MUSIQ): building a theory of context in healthcare quality improvement. BMJ Qual Saf. 2012;21(1):13-20.
110. Kilo CM. Improving care through Collaboration. Pediatrics. 1999;103(1):384-93.
111. Sifrim ZK, Barker PM, Mate KS. What gets published: the characteristics of quality improvement research articles from low- and middle-income countries. BMJ Qual Saf. 2012;21(5):423-31.

112. Homer CJ FP, Horvitz L, Peterson LE, et al. Impact of a quality improvement program on care and outcomes for children with asthma. Arch Pediatr Adolesc Med. 2005;159(5):464-9.
113. Bruce E.Landon IBW, Keith McInnes, Mary Beth Landrum, Lisa Hirschhorn, Peter Marsden, David Gustafson, Paul D.Cleary. Effects of a quality improvement collaborative on the outcomes of care of patients with HIV infection: The EQHIV study. Ann Intern Med. 2004;140(11):887-96.
114. Benedetti R, Flock B, Pedersen S, Ahern M. Improved clinical outcomes for fee-for-service physician practices participating in a diabetes care collaborative. Joint Commission journal on quality and safety. 2004;30(4):187-94.
115. Baier R, Gifford D, Lyder C, Schall M, Funston-Dillon D, Lewis J. Quality improvement for pressure ulcer care in the nursing home setting: the Northeast Pressure Ulcer Project. J AM Med Dir Assoc. 2003;4(6):291-301.
116. Mangione-Smith R, Schonlau M, Chan K, J K, Rosen M, Louis T, et al. Measuring the effectiveness of a collaborative for quailty improvement in paediatric asthma care: does implementing the chronic care mdoel improve processes and outcomes of care. Ambul Pediatr. 2005;5(2):75-82.
117. Asch SM, Baker DW, Keesey JW, Broder M, Schonlau M, Rosen M, et al. Does the collaborative model improve care for chronic heart failure? Med Care. 2005;43(7):667-75.
118. Pierce-Bulger M, Nighswander T. An Approach to reducing infant mortality using quality improvement principles. Quality management in health care. 2001;9(3):40.
119. Sodzi-Tettey S, Tum-Dwanso N, Mobisson-Etuk L, Macy L, Roessner J, Barker P. Lessons learned from Ghana's Proejct Fives Alive! A practical guide for designing and executing large-scale improvement initiatives. Cambridge, Massachussets: Institute for Healthcare Improvement, 2015.
120. O'Connor P, Rush W, Peterson J. Continuous quality imporvement can improve glycemic control of HMP patients with diabetes. Arch Fam Med. 1996;5:502-6.
121. Marlies Hulcher LS, Richard Grol. Collaboratives. Quest for Quality and Improved performance, The Health Foundation. 2009.
122. Howard D, Siminoff L, McBride V, al e. Does quality improvement work?Evaluation of the organ donation breakthrough collaborative. Health Serv Res. 2007;42:2160-73.
123. Pearson M, Wu S, Schaefer J, Bonomi A, Shortell S, Mendel P, et al. Assessing the implementation of the chronic care model in quality imporvement collaboratives. Health Serv Res. 2005;40(4):978-96.
124. Newton P, Davidson P, Halcomb E, Denniss A. Barriers and facilitators to the implementation of the collaborative method: reflections from a single site. Quality and Safety in health care. 2007;16:409-14.
125. Gollop R, Whitby E, Buchanan D, Ketley D. Influencing sceptical staff to become supporters of service improvement: a qualitative study of doctor's and managers' views. Quality and Safety in health care. 2004;13:108-14.
126. Wilson T, Berwick DM, Cleary PD. What do collaborative improvement projects do? Experience from seven countries. Joint Commission journal on quality and safety. 2003;29(2):85-93.

127. Mills PD, Weeks WB. Characteristics of successful quality improvement teams: lessons from five collaborative projects in the VHA. Joint Commission journal on quality and safety. 2004;30(3):152-62.
128. Ovretveit J, Gustafson D. Evaluation of quality improvement programmes. Qual Saf Health Care. 2002;11(3):270-5.
129. LeBrasseur R, Whissell R, Ojha A. Organizational learning, transformational leadership and implementation of continuous quality improvement in Canadian hospitals. Australian Journal of Management. 2002;27:141-62.
130. Weeks W, Mills P, Waldron J, Brown S, Spreoff T, Coulson L. A model for improving the quality and timeliness of compensation and pension examinations in the VA facilities. Journal of Healthcare management. 2003;48:252-61.
131. Bartlett M, Lloyd P, Berry G. Effective-TQM in accredited NSW hospitals: A secondary analysis of data. Journal of quality in clinical practice. 1997;17:177-85.
132. Nembhard I, Edmondson A. Making it safer: The effects of leader inclusiveness and professional status on psychological safety and improvement efforts in health care teams. Journal of organizational behaviour. 2006;27:941-66.
133. Shekelle PG. Why don't physicians enthusiastically support quality improvement programmes? Qual Saf Health Care. 2002;11(1):6.
134. Berner E, Baker C, Funkhouser E, Heudebert G, Allison J, Fargason Jr C, et al. Do local opinion leaders augment hospital quality improvement efforts? A randomized trial to promote adherence to unstable angina guidelines. Medical Care. 2003;41:420-31.
135. Wilkens R, London M. Relationship between climate, process and performance in continuous quality improvement groups. Journal of vocational behaviour. 2006;69:510-23.
136. Leatherman S, Sutherland K. Quality of care in the NHS of England. Bmj. 2004;328(7445):E288-90.
137. Berlowitz D, Young G, Hickey E, Saliba D, Mittman B, Czarnowski E, et al. Quality improvement implemntation in the nursing home. Health Serv Res. 2003;38:65-83.
138. Carman J, Shortell S, Foster R, Hughes E, Boerstler H, O'Brien J, et al. Key for successful implementation of total quality management in hospitals. Health Care Management Review. 1996;21:48-60.
139. Parker V, Wubenhorst W, Young G, Desai K, Charns M. Implementing quality improvement in hospitals: the role of leadership and culture. American Journal of Medical Quality 1999;14:64-9.
140. Shortell S, O'Brien J, Carman J, Foster R, Hughes E, Boerstler H, et al. Assessing the impact of continuous quality improvement/total quality management: Concpet versus implementation. Health Serv Res. 1995;30:377.
141. Wakefield B, Blegen M, Uden-Holman T, Vaughn T, Chrischilles E, Wakefield D. Organizational culture, continuous quality improvement and medication error reporting. American Journal of Medical Quality. 2001;16:128-34.
142. WHO. Quality of care: a process for making strategic choices in health systems. World Health Organization. 2006.
143. Wilson T, Holt T, Greenhalgh T. Complexity science: complexity and clinical care. Bmj. 2001;323(7314):685-8.

144. Lindenauer PK. Effects of quality improvement collaboratives. Bmj. 2008;336:1448-9.
145. Chassin M, O'Kane M. History of Quality Improvement Movement. In: Scott D, Kott A, editors. Towards improving the outcome of pregnancy III: March of Dimes; 2010.
146. Neuhauser D, Diaz M. Quality improvement research: are randomised trials necessary? Qual Saf Health Care. 2007;16(1):77-80.
147. Curran E, Benneyan J, Hood J. Controlling methicillin-resistant Staphylococcus aureus: A feedback approach using annotated statistical process control charts. Infect Control Hosp Epidemiol. 2002;23:13-8.
148. Wilcock P, Thomson R. Moderm measurement for a modern health service. Quality in health care : QHC. 2000;9:199-200.
149. Brown C, Lilford R. Evaluating service delivery interventions to enhance patient safety. Bmj. 2008;337:a2764.
150. Pronovost P, Bo-Linn G. Preventing patient harms through systems of care. JAMA. 2012;308(8):769-70.
151. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implementation science : IS. 2009;4:50.
152. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implementation science : IS. 2009;4(Additional File-3: CFIR constructs with short defintions):50.
153. R C. The importance of social theory for health promotion: from discussion to reflexivity. Health Prom Int. 1993;8(2):147-57.
154. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A. Synthesising qualitative and quantiative evidence: a review of possible methods. J Health Serv Res Policy. 2005;10(1):45-53.
155. Brown C, Hofer T, Johal A, Thomson R, Nicholl J, Franklin BD, et al. An epistemology of patient safety research: a framework for study design and interpretation. Part 4. One size does not fit all. Qual Saf Health Care. 2008;17(3):178-81.
156. Eccles M, Grimshaw J, Campbell J, Ramsay C. Research designs for studies evalauting the effectiveness of change and improvement strategies. Qual Saf Health Care. 2003;12:47-52.
157. Beard E, Lewis JJ, Copas A, Davey C, Osrin D, Baio G, et al. Stepped wedge randomised controlled trials: systematic review of studies published between 2010 and 2014. Trials. 2015;16(1):353.
158. Pawson R, N. T. Realistic Evaluation. Sage Publicatiion, London England. 1997.
159. Gustafson D, Quanbeck A, Robinson J, Ford II J, Pulvermacher A, French M, et al. Which elements of improvment collaboratives are most effective? A cluster-randomized trial. Addiction. 2013;108:1145-57.
160. Canadian Task Force on the Periodic Health Examination. The periodic health examination. Canadian Medical Association Journal. 1979;121(1193-1254).

161. Meldrum ML. A BRIEF HISTORY OF THE RANDOMIZED CONTROLLED TRIAL: From Oranges and Lemons to the Gold Standard. Hematology/Oncology Clinics of North America. 2000;14(4):745-60.
162. Investigators MS. Introduction of the Medical Emergency Team (MET) system: A cluster randomized controlled trial. Lancet. 2005(365):2091-7.
163. Shaw E, Ohman-Strickland P, Piasecki A, Hudson S, Ferrante J, McDaniel Jr R, et al. Effects of facilitated team meetings and learning collaboratives on colorectal cancer screening rates in primary care practices: A cluster randomized trial. Ann Fam Med. 2013;11(3):220-8.
164. Hawe P, Shiell A, Riley T. Complex interventions: how "out of control" can a randomised controlled trial be? Bmj. 2004;328(7455):1561-3.
165. Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, et al. Designing and evaluating complex interventions to improve health care. Bmj. 2007;334(7591):455-9.
166. Neuhauser D, Diaz M. Daniel:ising the Bible to teach quality improvement methods. Qual Saf Health Care. 2004(13):153-5.
167. Diaz M, Neuhauser D. Pasteur and parachutes: when statistical process control is better than a randomized controlled trial. Qual Saf Health Care. 2005;14(2):140-3.
168. Solodky C, Hegang C, Paul J, William K, Neuhauser D. Patients as partners in clinical research: A proposal for applying quality improvement methods to patient care. Medical Care. 1998;36(8):AS13-AS20.
169. Berwick DM. The science of improvement. JAMA. 2008;299(10):1182-84.
170. Brown C, Hofer T, Johal A, Thomson R, Nicholl J, Franklin BD, et al. An epistemology of patient safety research: a framework for study design and interpretation. Part 2. Study design. Qual Saf Health Care. 2008;17(3):163-9.
171. Leykum LK, Pugh JA, Lanham HJ, Harmon J, McDaniel RR, Jr. Implementation research design: integrating participatory action research into randomized controlled trials. Implementation science : IS. 2009;4:69.
172. Hargreaves JR, Copas AJ, Beard E, Osrin D, Lewis JJ, Davey C, et al. Five questions to consider before conducting a stepped wedge trial. Trials. 2015;16(1):350.
173. Pronovost P, Needham D, Berenholtz S, Sinopoli D, Chu H, Cosgrove S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. N Engl J Med. 2006;355(26):2725-32.
174. Perla RJ, Bradbury E, Gunther-Murphy C. Large-Scale Improvement Initiatives in Healthcare: A Scan of the Literature. Journal for Healthcare Quality. 2011:n/a-n/a.
175. Scott T, Mannion R, Davies HT, Marshall MN. Implementing culture change in health care: theory and practice. International journal for quality in health care : journal of the International Society for Quality in Health Care / ISQua. 2003;15(2):111-8.
176. Black R. The National Evaluation Platform Approach [26-Oct'15]. Available from: http://www.jhsph.edu/research/centers-and-institutes/institute-for-international-programs/current-projects/national-evaluation-platform/.
177. Kirkpatrick DL. Evaluating training programs; The four levels. 2nd ed. San Francisco, CA: Berrett-Koehler Publishers; 1998.

178. Barr H, Hammick M, Koppel I, Reeves S. Evaluating interprofessional education:Two systematic reviews for health and social care. British Educational Research Journal. 1999;25(4):533-43.
179. Dixon-Woods M, McNicol S, Martin G. Ten challenges in improving quality in healthcare: lessons from the Health Foundation's programme evaluations and relevant literature. BMJ Qual Saf. 2012(21):876-84.
180. Langley A, Denis JL. Beyond evidence: the micropolitics of improvement. BMJ Qual Saf. 2011;20 Suppl 1:i43-6.
181. Rogers E. Diffusion of Innovations. New York: Free Press; 1995.
182. Greenhalgh T, Robert G, MacFarlane F, Bate P, Kyriakidou O. Diffusion of Innovations in service organizations: Systematic review and reccomendations. The Milbank Quarterly. 2004;82(4):581-629.
183. May C. Towards a general theory of implementation. Implementation science : IS. 2013;8:18.
184. Kippax S. Sexual health interventions are unsuitable for experimental evaluation. in Effective sexual health interventions: issues in experimental evaluation. 2003;Chapter 2:17-34.
185. Pronovost P, Miller M, Wachter R. The GAAP in quality measurement and reporting. JAMA. 2007;298(15):1800-2.
186. Auerbach AD, Landefeld CS, Shojania KG. The tension between needing to improve care and knowing how to do it. N Engl J Med. 2007;357(6):608-13.
187. World Health Organization. International Classification of Diseases, Tenth Revision (ICD-10). 2007.
188. Gabrysch S, Civitelli G, Edmond KM, Mathai M, Ali M, Bhutta ZA, et al. New signal functions to measure the ability of health facilities to provide routine and emergency newborn care. PLoS Med. 2012;9(11):e1001340.
189. Vergnano S, Fottrell E, Osrin D, Kazembe PN, Mwansambo C, Manandhar D, et al. Adaptation of a probabilistic method (InterVA) of verbal autopsy to improve the interpretation of cause of stillbirth and neonatal death in Malawi, Nepal and Zimbabwe. Population Health Metrics. 2011;9(48).
190. Perla RJ, Provost LP, Murray SK. The run chart: a simple analytical tool for learning from variation in healthcare processes. BMJ Qual Saf. 2011;20(1):46-51.
191. Kanyuka M, Ndawala J, Mleme T, Chisesa L, Makwemba M, Amouzou A, et al. Malawi and Millennium Development Goal 4: Early Adopter; Early Achiever. Lancet Global Health (under peer review). 2015.
192. Reinersten JL, Gosfield A, Rupp W, JW W. Engaging Physicians in a Shared Quality Agenda. Cambridge, Massachusetts: Institute for Healthcare Improvement, 2007.
193. Hayes R, Moulton L. Cluster Randomized Trials2009.
194. Pawson R, Greenhalgh J, Brennan C, Glidewell E. Do reviews of healthcare interventions teach us how to improve healthcare systems? Soc Sci Med. 2014;114:129-37.
195. Habicht JP, Victora CG, Vaughan JP. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. Int J Epidemiol. 1999;28(1):10-8.
196. Grol R, Baker R, Moss F. Quality improvement research: understanding the science of change in health care. Qual Saf Health Care. 2002;11(2):110-1.
197. Chen H, Rossi PH. Theory driven evaluation. Sage Publication. 1990.

198. Blamey.A, Mackenzie M. Theories of Change and Realistic Evaluation: Peas in a Pod or Apples and Oranges? Evaluation. 2007;13:439-56.
199. Mason.P., M B. Constructing theories of change. Evaluation. 2007;13(2):151-70.
200. Marchal B, Dedzo M, Kegels G. A realist evaluation of the management of a well-performing regional hospital in Ghana. BMC health services research. 2010;10:24.
201. Wong G, Westhorp G, R P, Greenhalgh T. Realist Synthesis. RAMESES Training materials. 2013:1-55.
202. Pawson R. The Science of Evaluation. London, UK: SAGE; 2013. 210 p.
203. Bonner L. Using Theory-Based Evaluation to Build Evidence-Based Health and Social Care Policy and Practice. Critical Public Health. 2003;13(1):77-92.
204. Van Belle SB, Marchal B, Dubourg D, Kegels G. How to develop a theory-driven evaluation design? Lessons learned from an adolescent sexual and reproductive health programme in West Africa. BMC Public Health. 2010;10:741.
205. Powell BJ, MCMillen JC, Proctor EK, Carpenter CR, Griffey RT, Bunger AC, et al. A compilation of strategies for implementing clinical innovations in health and mental health. Medical Care Research and Review. 2012;69(2):123-57.
206. Gilson L. Part 1: Introduction to health policy and systems research. In: Gilson L, editor. Health policy and systems research: a methodology reader. WHO, Geneva: World Health Organization; 2012.
207. P. Rogers, S F. Pruposeful Program Theory: Effective useof Theories of Change and Logic Models. . Jossey Bass. 2011.
208. Smits H, LeCompte R, Kotagal U, van der Broke N, Flatman D, Nambiar B, et al. Proposal to The Health Foundation for Improvement of the Quailty of maternal and neonatal care in Malawi. The Health Foundation, 2005 October, 2005. Report No.
209. Institute for Healthcare Improvement. The Breakthrough Series: IHI's Collborative Model for Achieveing Breakthrough Improvement. IHI Innovation Series White Paper. 2003.
210. DfID. EHRP Evaluation- Final Report. 2010:1-86.
211. McCannon CJ, Berwick DM, Massoud MR. The science of large-scale change in global health. JAMA. 2007;298(16):1937-9.
212. Graham WJ, Bell J, Bullough C. Can skilled attendacne at delivery reduce maternal mortality in developing countries. Studies Health Serv Organ Policy. 2001;17:97-130.
213. Bayley O, Colbourn T, Nambiar B, Costello A, Kachale F, Meguid T, et al. Knowledge and perceptions of quality of obstetric and newborn care of local health providers: a cross-sectional study in three districts in Malawi. Malawi Med J. 2013;25(4):105-8.
214. Nsona H, Mtimuni A, Daelmans B, Callaghan-Kori JA, Gilroy L, Mgalula L, et al. Scaling up Integrated Community Case Management of Childhood Illness: Update from Malawi. Am J Trop Med Hyg. 2012;87(Suppl 5):54-60.
215. Godlonton S, Okeke EN. Does a ban on informal health providers save lives? Evidence from Malawi. Journal of Development Economics. 2016;118:112-32.
216. Leigh B, Mwale TG, Lazaro D, Lunguzi J. Emergency obstetric care: How do we stand in Malawi? International Journal of Gynecology and Obstetrics.101(1):107-11.

217. Greenhalgh T, Humphrey C, Hughes J, Macfarlane F, Butler C, Pawson RAY. How Do You Modernize a Health Service? A Realist Evaluation of Whole-Scale Transformation in London. Milbank Quarterly. 2009;87(2):391-416.
218. Allison L-S, Donald S, Dale M N, Colantuoni E, Morlock LL, Pronovost PJ. Impact of a statewide intensive care unit quality improvement initiative on hospital mortality and length of stay: retrospective comprative analysis. Bmj. 2011(342: d219).
219. Ferguson BS. Profits and the Hospital Sector: What does the literature really say? Health Policy Working Paper; Atlantic Institute for Market Studies. 2002:1-23.
220. Astbury B, Leeuw FL. Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation. American Journal of Evaluation. 2010;31(3):363-81.
221. Tilley N. Theory Based Evaluation: EVALSED; 2012.
222. Chen H-T. The conceptual framework of the theory driven perspective. Evaluation and Program Planning. 1989;12:391-6.
223. Rickles D, Hawe P, Shiell A. A simple guide to chaos and complexity. J Epidemiol Community Health. 2007;61(11):933-7.
224. Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. Bmj. 2008;336(7656):1281-3.
225. Alemi F, Safaie F, Neuhauser D. A survey of 92 quality improvement projects. JT Comm J Qual Improve. 2001;27:619-32.
226. Dixon-Woods M, Leslie M, Tarrant C, Bion J. Explaining Matching Michigan: an ethnographic study of a patient safety program. Implementation science : IS. 2013;8:70-.
227. agyeman-Duah J, Theurer A, Munthali C, Alide N, Neuhann F. Understanding the barriers to setting up a healthcare quality improvement process in resoruce-limited settings: a situational analysis at the Medical Department of Kamuzu Central Hospital in Lilongwe, Malawi. BMC health services research. 2014;14:1-10.
228. de Savigny D, Adam T. Systems thinking for health systems strengthening. In: WHO, editor. Geneva2009. p. 1-112.
229. Stame N. Theory based evaluation and types of complexity. Evaluation. 2004;10(1):58-76.
230. Massoud R, Karen A, Reinke J, Franco L, BOrnstein T, Knebel E, et al. A Modern Paradigm for Improving Healthcare Quality. Quality Assurance Project, USAID. 2002:1-90.
231. Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging research and practice: models for dissemination and implementation research. American journal of preventive medicine. 2012;43(3):337-50.
232. Grol RP, Bosch MC, Hulscher ME, Eccles MP, Wensing M. Planning and studying improvement in patient care: the use of theoretical perspectives. Milbank Q. 2007;85(1):93-138.
233. De Silva MJ, Breuer E, Lee L, Asher L, Chowdhary N, Lund C, et al. Theory of Change: a theory-driven approach to enhance the Medical Research Council's framework for complex interventions. Trials. 2014;15:267.
234. Ridde V. Per diems undermine health interventions, systems and research in Africa: burying our heads in the sand. Trop Med Int Health. 2010.

235. Glouberman S, Zimmerman B. Complicated and complex systems. What would successful reform of Medicare look like? Commission on the Future of Health Care in Canada. 2002;Discussion Paper 8.
236. Luke DA, Harris JK. Network analysis in Public Health: History, Methods and Applications. Annu Rev Public Health. 2007;28:69-93.
237. Lombarts MJ, Rupp I, Vallejo P, Klazinga NS, Sunol R. Differentiating between hospitals according to the "maturity" of quality improvement systems: a new classification scheme in a sample of European hospitals. Qual Saf Health Care. 2009;18 Suppl 1:i38-43.
238. Bergstrom A, Skeen S, Duc DM, Blandon EZ, Estabrooks C, Gustavsson P, et al. Health system context and implementation of evidence-based practices-development and validation of the Context Assessment for Community Health (COACH) tool for low- and middle-income settings. Implementation science : IS. 2015;10(1):120.
239. Leeuw F. Report on Theory-Based Evaluation (TBE), Prepared for the Evalsed Handbook on Impact Evaluation. DG Regional Policy, Brussels. in press.
240. Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review--a new method of systematic review designed for complex policy interventions. J Health Serv Res Policy. 2005;10 Suppl 1:21-34.
241. RAMESES Archives. https://www.jiscmail.ac.uk/cgi-bin/webadmin?A1=ind1506&L=RAMESES 2015. Available from: https://www.jiscmail.ac.uk/cgi-bin/webadmin?A1=ind1506&L=RAMESES.
242. Feinstein O. Realistic Evaluation (book review). Evaluation. 1998;4(2):243-6.
243. Coryn LSC, Noakes LA, Westine CD, Scroter DC. A Systematic Review of Theory-Driven Evaluation Practice from 1990 to 2009. American Journal of Evaluation. 2011;32(2):199-226.
244. Patton MQ. Developmental Evaluation: Applying complexity concepts to enhance innovation and use. New York: NY: Guilford; 2010.
245. Leischow SJ, Milstein B. Systems thinking and modeling for public health practice. Am J Public Health. 2006;96(3):403-5.
246. Bryce J, Victora C, Boerma T, Peters D, Black RE. Evaluating the scale-up for maternal and child survival: A common framework. Working paper IIP-Johns Hopkins Bloomberg School of Public Health. 2010.
247. Victora CG, Black RE, Boerma JT, Bryce J. Measuring impact in the Millennium Development Goal era and beyond: a new approach to large-scale effectiveness evaluations. Lancet. 2011;377(9759):85-95.
248. Rogers PJ. Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. Evaluation. 2008;14(1):29-48.
249. Greenhalgh T, Clinch M, Afsar N, Choudhury Y, Sudra R, Campbell-Richards D, et al. Socio-cultural influences on the behaviour of South Asian women with diabetes in pregnancy: qualitative study using a multi-level theoretical approach. BMC Med. 2015;13:120.
250. Hawkins A. The case for experimental design in realist evaluation. Learning Communities International Journal of Learning in Social Contexts 2014;14(Special Issue: Evaluation):46-59.
251. Coryn CLS. Program theory-driven evaluation science. American Journal of Evaluation. 2008;29:215-20.

252. McMullen H, Griffiths C, Leber W, Greenhalgh T. Explaining high and low performers in complex intervention trials:a new model based on diffusionof innovations theory. Trials. 2015;16:242.
253. Global Health Workforce Alliance. A universal truth: no health without a workforce. World Health Organization. 2014.
254. Bonell C, Fletcher A, Morton M, Lorenc T, Moore L. Methods don't make assumptions, researchers do: a response to Marchal et al. Soc Sci Med. 2013;94:81-2.
255. Marchal B, Westhorp G, Wong G, Van Belle S, Greenhalgh T, Kegels G, et al. Realist RCTs of complex interventions - an oxymoron. Soc Sci Med. 2013;94:124-8.
256. Black RE. National Evaluation Platform approach. Available from: http://www.jhsph.edu/research/centers-and-institutes/institute-for-international-programs/current-projects/national-evaluation-platform/.
257. Baio G, Copas A, Ambler G, Hargreaves J, Beard E, Omar RZ. Sample size calculations for a stepped wedge trial. Trials. 2015;16:354.
258. Grove J, Claeson M, Bryce J, Amouzou A, Boerma T, Waiswa P, et al. Maternal, newborn, and child health and the Sustainable Development Goals—a call for sustained and improved measurement. The Lancet.386(10003):1511-4.
259. Gaffar A. Presentation at the Health Systems Research Conference, Cape Town. South Africa. 2014.
260. Don de Savigny, Taghreed A. Systems thinking for health systems strengthening. Alliance for Health Policy and Systems Research. 2009.
261. Baio G, Copas A, Ambler G, Hargreaves J, Beard E, Omar RZ. Sample size calculation for a stepped wedge trial. Trials. 2015;16(1):354.
262. Souza JP, Gulmezoglu AM, Vogel J, Carroli G, Lumbiganon P, Qureshi Z, et al. Moving beyond essential interventions for reduction of maternal mortality (the WHO Multicountry Survey on Maternal and Newborn Health): a cross-sectional study. Lancet. 2013;381(9879):1747-55.
263. Donabedian A. The definition of quality and approaches to its assessment. Ann Arbor, Michigan: Health Administration Press; 1980. 163 ; 23 cm. p.
264. Davidoff F, Dixon-Woods M, Leviton L, Michie S. Demystifying theory and its use in improvement. BMJ Qual Saf. 2015;24(3):228-38.
265. Popper K. The Logic of Scientific Discovery. 5th edn ed. London: Routledge; 1992.

Appendix 1 Training Summary

Departmental Seminar Attendance - Per Term	
Socio-economic determinants of household food security and womenâ€™s dietary diversity in rural Bangladesh: a cross-sectional study-Helen Fry	
1	15/01/14
Conference - Attendance	
Maimwana Dissemination conference, Lilongwe Malawi	
2	11/10/13
Intensive Course in Epidemiology and Medical Statistics (ICEMS)	
Organiser: LSHTM	
10	01/07/13
Literature Searching using PubMed	
Literature search on quality improvements in maternal and child health in Africa	
1	01/06/13
MPhil to PhD Upgrade Session for Research Students	
Upgrade	
1	06/05/13
Co-ordinating / Convening a Conference or Workshop	
Coordinating a workshop on UCL wide quality improvement collaborations	
5	01/06/12
Report/Paper Preparation	
The MaiKhanda Report submitted to The Health Foundation	
2	20/01/12
Using Word 2013 for Dissertations/Theses Using Word for Dissertations/Theses - 2009/10	
Once you have completed your training, please log your points here.	
0	1
Introduction to Survival Analysis	
1	23/04/10
Introduction to Logistic Regression	
1	21/04/10
Introduction to Regression Analysis	
1	16/04/10
Introduction to Regression Analysis	
1	15/04/10
Research Student Induction Session	
1	15/04/10
Conference - Attendance	
Attended the International Conference on Quality and Safety in HealthCare in Berlin in 2009. THis was a 4 day conference.	
Made presentations at the DevCo meeting which was held earlier on the 16th of March, 2009	
2	17/03/09
Fieldwork - 6-12 Months	
Based in Malawi carrying out field based research on maternal and neonatal health.	
20	02/01/07

Appendix 2 Summary of individual studies

2.1 Baseline survey 2006

Study Title	Baseline survey 2006
Study objective	To identify and assess the capacity of existing health facilities and document the level of functioning of facilities for delivery of MNH services prior to introduction of MaiKhanda interventions
Study design	Cross sectional survey
Methods/Tools	All District HMIS[xvii] Officers were approached to obtain general district information such as district population and list of health facilities in the district. A Health Facility Rapid Assessment Tool (RAT) was used to collect data from the health facilities.
Study period	June-August, 2006. Data was collected over a 3 month period April, May and June 2006.
Sample	All hospitals and HCs in the three target districts of Kasungu, Lilongwe ad Salima, which provide maternity services, regardless, of managing authority, were included in the survey.
Data management	All data from the data collection forms were double entered by 2 data entry clerks, using SPSS version 12.
Analysis plan	Data was analysed and tabulated using SPSS v.12 and reported by district.

[xvii] *HMIS-Health Management Information System. Division of the Ministry of Health responsible for Monitoring and Evaluation*

2.2 Baseline evaluation 2007

Study Title	Baseline evaluation 2007
Study objective	To provide an overview of the programme implementation at the beginning of the pre-intervention period
Study design	Retrospective evaluation
Methods/Tools	Tools used during this evaluation included document searches, interviews, questionnaires and observations. Documents were sourced from THFC archives and from facilities. The documents included monthly reports, QI tools such as checklists employed by the facilities, and THFC charter. Anonymous questionnaires were distributed to all QI team members in the nine CEmOC facilities. These questionnaires were followed up with semi-structured interviews with representative QI team members in MaiKhanda. In-depth interviews were also conducted with key staff involved in the QI implementation
Study period	June-August, 2007. Data was collected over a 3 month period
Sample	CEmOC facilities in the three target districts of Kasungu, Lilongwe ad Salima, where QI interventions were in operation over the past 12 months
Data management	All data was collected by a Research Assistant from IGH.
Analysis plan	Data was analysed and tabulated using qualitative research methods.

2.3 Health facility resources survey

Objective	The objective was to understand the context in which QI interventions were being implemented
Study design	Cross-sectional survey
Methods/Tools	Survey questionnaire
Study period	First half of 2010
Sample	69 health centres: 37, 18 & 14 in LL, KU & SA
Data collection	The questionnaire was completed by the MEOs during their routine field visit to the health facilities.
Key variables	Staff availability, physical infrastructure, material inventories, transport and communication
Analysis plan	Analysis will include a comparison of resource availability in intervention and control areas.
Reference in MaiKhanda Evaluation Report(186)	§2.2.5.1.1 (p.43) §4.2.5.1.1 (p.174) Tools: Appendix 10: Health Facility Surveillance Form (p.280) Appendix 11: Health Centre Resource Survey (p.281)

2.4 Provider knowledge survey

Objective	To assess provider's knowledge of best practice as well as their perception of the care they provide
Study design	Cross-sectional survey
Methods/Tools	Self completed questionnaire
Study period	December 2009- March 2010
Sample size	A purposive sample of 52 clinical staff comprising of different cadres across Lilongwe, Kasungu and Salima districts and CEmOC facilities were interviewed
Training	
Data management	The self-completed forms were complted by the staff in the facility with MEOs readily available to clarify any queries. The data was completed by the MEOs and returned to offices to be entered into an excel database. Data was analysed using SPSS (V14.0)
Key variables	Questions related to emergency obstetric and newborn care, monitoring and routine labour
Analysis plan	Comparison by staff category, between districts and between health centres and CEmOC facilities
Reference in MaiKhanda Evaluation Report(186)	§2.2.5.1.3 (p.44) §4.2.5.1.3 (p.179) Tool: Appendix 14: Provider Knowledge Survey (p.289)

2.5 Staff Psychology survey

Objective	To understand the factors influencing staff motivation
Study design	Cross-sectional survey
Methods/Tools	Mixed methods: self administered questionnaire & interviews
Study period	Data was collected over a period of 5 months from January –May 2008.
Sample	132 staff members across 57 health facilities and 15 in-depth interviews. It was not a random sample-more based on availability of staff during the MEO field visit-purposive sample
Data management	Forms were made available to the staff during their monthly field visits. These were completed by staff in their respective facilities and returned to MEOs during their next visit. The data was entered in SPSS (v14.0) The interviews were done in Chichewa or English depending on the preference of the interviewee. The interviews were recorded using digital recorders and data transferred to the computer and transcribed by MaiKhanda data coordinator.
Key variables	19 set question related to psychology
Analysis plan	Analysis will include an analysis of the factors affecting staff motivation Transcripts were analysed by a qualitative researcher using standard qualitative techniques
Reference in MaiKhanda Evaluation Report(186)	§2.2.5.1.2 (p.43) §4.2.5.1.2. (p.178) Tools: Appendix 12: Staff Psychology Survey (p.284) Appendix 13: Interview Guidelines (p.287)

2.6 CEmOC survey

Objective	To assess the QI activities being carried out in CEmOC facilities and to gauge the understanding and general impression regarding model for improvement among QI team members.
Study design	Cross-sectional survey
Methods/Tools	Self completed questionnaire
Study period	August 2008
Sample size	All of 27 QI team members from the CEmOC facilities were approached of which 25 responded to the questionnaire.
Data management	The tool was distributed and collected by the Senior Monitoring and Evaluation (M&E) Officer for Lilongwe, with the help of Assistant QI Officers. The data was compiled by the MEOs and returned to offices to be entered into an excel database. Data was analysed in excel.
Reference in MaiKhanda Evaluation Report(186)	This was not included in the MaiKhanda evaluation report. But was part of the evaluation report submitted to MaiKhanda as part of the Quarterly Progress Report, May 2009

2.7 St.Gabriel case study: QI approaches to newborn resuscitation

Background

In addition to the various sub-studies that were done to enhance the knowledge regarding QI interventions, a case study of the neonatal resuscitation improvement exercise at S.Gabriels was also conducted to add to the body of knowledge regarding implementation of QI interventions in Malawi.

Introduction

St.Gabriel is a missionary hospital, established in 1959 and managed by the Christian Health Association of Malawi (CHAM). It is located approximately 53 kms to the west of Lilongwe district in Chileka health area and provides health services to a population of more than 200,000 including some of the neighbouring population of Mchinji district. It is one of 8 health facilities providing comprehensive emergency obstetrics (CEmOC) services within Lilongwe district.

St.Gabriel has an average of 235 deliveries per month. The maternity wing consists of a four bedded labour ward and postnatal ward. It has two operation theatres with one of them equipped with a neonatal resuscitation unit. The other resuscitation unit is in the labour ward. The labour ward has 1 Clinical Officer, 6 NMTs (Nurse Midwife Technicians), and 5 patient attendants.

MaiKhanda was working with St. Gabriel's hospital since 2007 with a 10 member QI (Quality Improvement) team there.

Problem statement

In 2009, the NCFR (neonatal case fatality rate) at St. Gabriel was 22 per 1000 live births, with asphyxia being one of the major contributors (27%).

In a survey on provider knowledge conducted by MaiKhanda in 2010 amongst 52 health facility staff in Kasungu, Lilongwe and Salima, the biggest gaps in knowledge in neonatal care were in the management of a baby who does not breathe spontaneously (only 35% answered this correctly) and management of a baby with low Apgar scores at 1 minute: 23% of respondents gave potentially life-threatening responses, including 8% who said no action was required and 4% who said they did not know what to do.

The Principal Nursing Officer at St. Gabriel hospital had similar observations as well

"Last year when we reviewed our neonatal deaths, those occurring from asphyxia-there were a lot of them. From observation of our nurses doing resuscitation, we realised the nurses did not have the technique and attempts at resuscitation resulted in failure".

Intervention

MaiKhanda was working with the QI team at St. Gabriel since 2009 to reduce neonatal deaths due to asphyxia. The primary objective was to improve the skills and competency of the delivery team so that they are able to conduct neonatal resuscitation within one minute of delivering a baby (referred to as the 'golden minute').

Staff at St. Gabriel were initially trained by a team from the US using the - "Helping Babies Breathe" model developed by the American Association of Paediatrics. This was followed by a refresher on-site training (also referred to as in-situ training), which initially included all staff, followed by a fortnightly training (Thursdays) for the maternity staff, not lasting more than 30 minutes and supervised by the clinical officer. A resuscitation drill kit (including the mannequin) ()Figure 25) was available in the labour ward at all times for the staff to repeat the drill at frequent intervals. MaiKhanda's main role was in facilitating the refresher training.

Figure 25: Neonatal resuscitation demonstration kit

During a neonatal resuscitation drill every nurse demonstrated their resuscitation skills on a mannequin and each drill got timed. The target was to complete the resuscitation drill correctly in one minute (as this is the amount of time they would have in a real life situation). The nurses were given a score at the end of each drill for their skill. After they completed their drill, the nurses themselves would point out what they did better and what they could have done better. The team would also give their comments and report how the resuscitation could have been done better.

As part of the resuscitation protocol, for every birth with low Apgar score staff in the labour ward would give antibiotic prophylaxis.

Another important feature of the intervention was that the patient attendants also received training on how to resuscitate the babies so as to improve teamwork and initiate prompt management in case of labour emergencies.

"...for instance, sometimes you have cases where the mother goes into post partum haemorrhage after delivery and the baby is asphyxiated at the same time. This puts the nurse in a dilemma as whom to attend first. Having the patient attendants trained in resuscitation, helps the nurse on duty because then she can focus on one thing…"

Results:

The staff at St. Gabriel hospital mentioned that neonatal resuscitation drills have had a positive effect because it was improving their skills and has also improved their timing in helping babies to breathe.

"My timing has improved; I can resuscitate a baby within a maximum of 3 minutes whilst before I would take 10 -11 minutes and would still be struggling

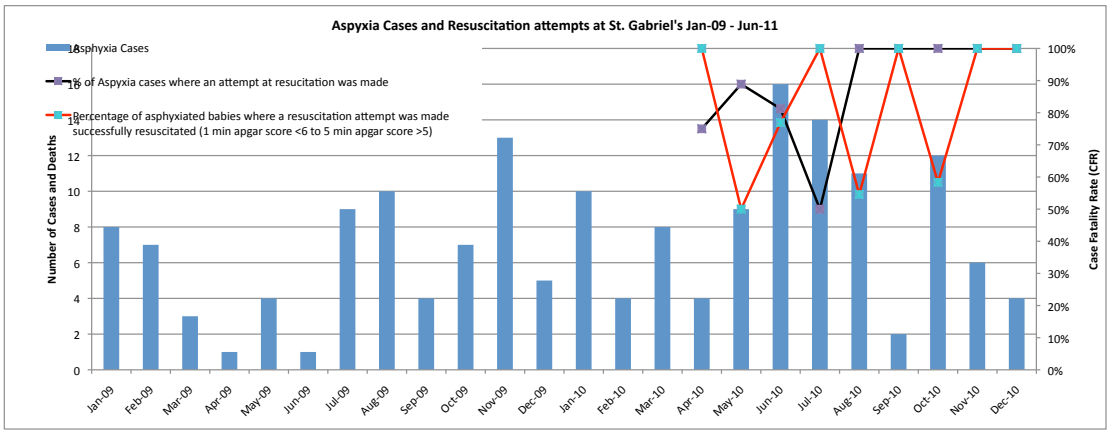

Figure 26: St.Gabriel hospital-Asphyxia cases and neonatal resuscitation

In the last quarter of 2010,100% of the asphyxiated babies received resuscitation, and most of them were successfully resuscitated (Figure 26).

This in turn has had a positive effect on staff morale and confidence in managing asphyxiated babies.

"… just coming straight from school and this my first posting. Before I was looking at neonatal resuscitation as something very difficult for a baby who was not breathing; now I don't have any fears even if I am left alone in labour ward to resuscitate a baby whereas before I had a fear that I would kill the baby and the mother would blame me; I used to be really anxious but now I am confident.."

The level of confidence and self efficacy in dealing with neonatal resuscitation is evident from the excerpt provided below.

Asphyxia case management at St. Gabriel Hospital - excerpts from an interview with nursing staff in labour ward.

"The last case of asphyxia that I had was yesterday. There was a mother who came with APH [ante-partum haemorrhage] *at 7 months. It was decided to take this mother for C-section* [Caesarean section]*… as soon as I was informed of the case being a pre-term, I prepared the resuscitation tray. Then went to collect the baby from the OT (Operation Theatre)and as soon as the baby was delivered rushed to labour ward with the baby and did my resuscitation.*

After doing everything baby cried for 30 seconds and then it stopped. There was no secretion and airway was clear so tried to commence O_2 therapy. Meanwhile the baby was blue but the heartbeat was still there. I tried the ambu bag and then the baby cried. I commenced O_2 [Oxygen therapy] *and started antibiotics and left baby on same O_2 therapy…"*

Whilst it is clear that the intervention has had an effect on improving staff skills and competency which resulted in a majority of the asphyxia cases being resuscitated, the study explored if there was an effect on asphyxia related mortality and neonatal case fatality rate at St. Gabriel's hospital.

As Figure 27 shows, the asphyxia case fatality rate (CFR) at St. Gabriel's seems to have gone down, despite a higher incidence of asphyxia cases from Jun-10 to Aug-10 (the higher incidence could be attributed to a better recording of cases).

It is notable that there were no deaths from asphyxia between Sep-10 and Dec-10. The zero deaths in these 4 months match with 100% of asphyxiated babies having a resuscitation attempt.

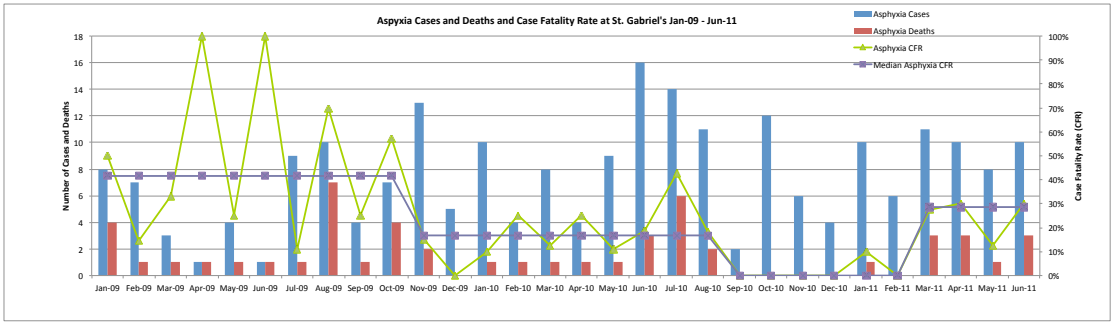

Figure 27: Data on Asphyxia at St. Gabriel's hospital

In order to ensure sustainability of the resuscitation intervention the staff at St.Gabriel felt that it was important to sensitize mothers to identify the danger signs during pregnancy so that they present themselves early at the hospital. Secondly to avoid eventuality because of staff attrition, there was a need to keep practising neonatal resuscitation regularly. Keeping the staff skills updated as frequently as possible was thought to be important for sustainability, as one of the interviewees stated:

"people have a recess and go backwards; still need to keep encouraging teams to keep practising".

The staff at St. Gabriel identified MaiKhanda to have an influential and complementary role to play in the intervention by helping identify the problem area, develop plans and implement them

"…sometime you are doing things blindly but the way MaiKhanda has helped is to look back on data, analyse data and try to learn from data and implement what can make the situation better".

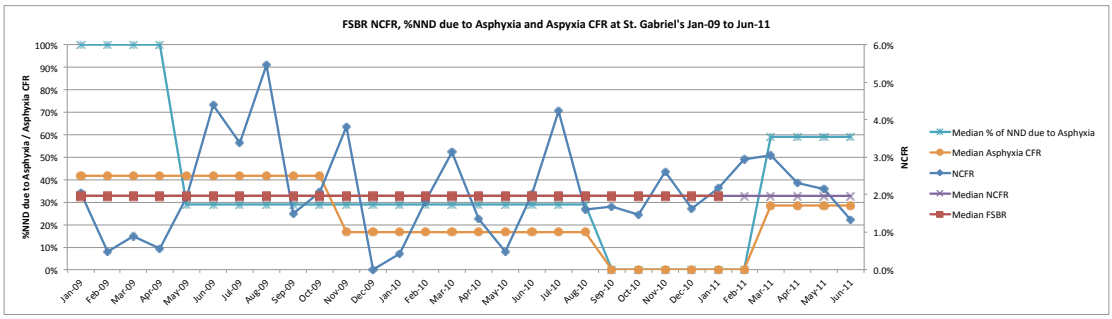

Figure 28: St.Gabriel's hospital asphyxia and neonatal deaths

Looking at how the total neonatal deaths relate to asphyxia deaths at St. Gabriel's (Figure 28: blue and purple lines vs. orange and turquoise lines;

medians change after 6 points below current median) it is apparent that the falls in the asphyxia CFR in late 2010 do not seem to be related to the total burden of NND at St. Gabriel, which has an unchanging median throughout the whole period (Jan-09 to Jun-11).

Discussion

By a global estimate, 3%-6% of all babies born in a facility require basic neonatal resuscitation[xviii]. By this estimate, 7 to 14 children born at St. Gabriel every month, will require resuscitation. Training providers in neonatal resuscitation in health facilities has the potential to reduce deaths due to asphyxia in full term babies.

At St. Gabriel, on average, staff attend 235 deliveries and resuscitate about 7 to 8 newborns per month. As these numbers are small and infrequent, it poses a challenge in maintaining a high standard of resuscitation knowledge and skills of the staff at all times. Refresher training is recommended every 6 months to prevent loss of skills. A conventional 5-day classroom training for 128 staff costs US $8,160 per course[xix]. Refresher trainings may cost a bit less than a 5 day classroom training, but would still pose a significant financial burden for organizations if conducted every 6 months.

St.Gabriel's neonatal resuscitation training method, consisting of on-site regular resuscitation drills at the facility under supervision, provides a low-cost alternative to maintaining the competency and skills of clinical staff. It also has the added advantages of staff remaining in the facility during training, and the training occurring at a much higher frequency of every 2 weeks as opposed to every 6 months. This improves the dose and duration of the intervention and

---

[xviii] *S Wall et al. Neonatal resuscitation in low-resource settings: What, who and how to overcome challenges to scale up?International Journal of Obstetrics and Gynaecology 2009 (107)-S47-S64*

[xix] *Based on estimates from Vidal et al study in Brazil covering 4 hospital and 128 staff. [Vidal et al. Comparison of two training strategies for essential newborn care in Brazil. Bulletin of the World Health Organization, 2001; 79(11)]*

depending on participation of team leaders & MaiKhanda staff, can also increase the intensity of the intervention.

Using a health facility perspective, a rough estimate puts the cost of conducting the neonatal resuscitation drill at St. Gabriel at US $0.22 per baby delivered. This includes the cost of initial training, the cost of four Laerdal neonatal mannequins and refresher training costs over a five year period (but does not take into account MaiKhanda's follow up cost). The costs for a similar training in Indonesia was US$0.16 per baby delivered[xviii].

There are other actors as well involved in improving resuscitation efforts at St. Gabriel (such as a US agency providing the initial training on neonatal resuscitation to all staff) and this contextual factor needs to be acknowledged in attribution programme success.

One of the concerns emerging from the success of the improvement efforts in resuscitation, is whether this will lead to increased morbidity and mortality of infants surviving resuscitation, leaving overall burdens unchanged. The burden of subsequent neonatal mortality and morbidity in the community also needs to be assessed.

Summary

An in-house resuscitation drill for the staff, conducted every 2 weeks, using a resuscitation protocol and a scoring system on staff performance was implemented at St. Gabriel since 2009 to tackle the high rate of asphyxia related neonatal deaths.

The exercise was a team effort, involved all staff, including patient attendants, in the labour ward, with the target of completing the neonatal resuscitation within the 'golden minute'.

Training tools such as a resuscitation mannequin were provided to the staff and were available in the labour ward at all times for staff to practice, in addition to the fortnightly drills.

A direct outcome of the intervention was improvements in the competency and skills of the nursing staff in labour ward at St. Gabriel hospital to conduct neonatal resuscitation. Their target was to complete the drill within the first minute of birth-and the staff felt competent enough to achieve that. The process had a positive effect on their morale and confidence as well.

However, despite these improvements, there were no changes to the overall neonatal case fatality rates at St.Gabriel

Sensitizing mothers on the danger signs of pregnancy and establishing links with the community for their continued support were considered crucial in supporting the neonatal efforts at the facility.